DOCUMENT RESUME

ED 366 617 TM 020 989

TITLE	Results of the 1992 Primary Assessment Program in Urban Schools.
INSTITUTION	Northern Territory Dept. of Education, Darwin (Australia).
REPORT NO	ISBN-0-7245-2637-4
PUB DATE	93
NOTE	64p.
PUB TYPE	Statistical Data (110) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Academic Achievement; Achievement Gains; Comparative Analysis; *Educational Assessment; Elementary Education; Elementary School Students; Foreign Countries; Graphs; *Indigenous Populations; *Mathematics Achievement; Program Evaluation; *Reading Achievement; Scores; Sex Differences; Tables (Data); Test Construction; *Test Results; Test Validity; *Urban Schools; Writing Achievement
IDENTIFIERS	Australia (Northern Territory); *Primary Assessment Program (Australia); Review Panels

ABSTRACT

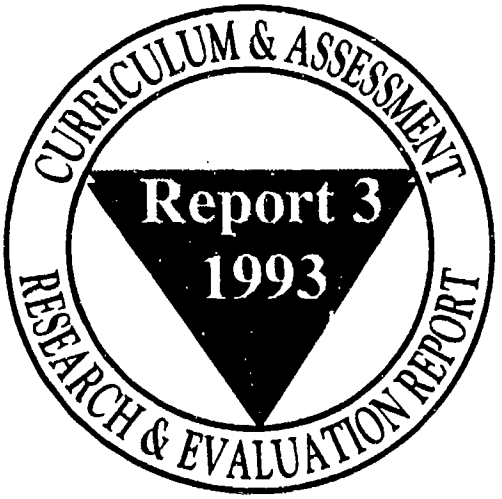
This report describes the aggregated reading and mathematics performances of students in years 5 and 7 in urban Northern Territory (Australia) schools on the Primary Assessment Program (PAP) tests. The aggregated results are encouraging, with apparent gains in some reading tests over the preceding year, and no evidence of decline in the others. Reading-test results indicate that the majority of students in years 5 and 7 are well able to read and comprehend texts considered suitable to their levels. There is evidence that boys marginally outperformed girls in the measurement strand of mathematics. On all reading tests, girls outperformed boys. Results of the mathematics tests indicate that students find these somewhat more difficult than the reading tests, with the average mark for both years at approximately 55 percent. Test results of urban self-identified Aboriginal students indicate that, on average, their achievement levels are well below those of non-Aboriginal students. Fourteen tables, 11 figures, and 42 graphs illustrate student achievement results. Six appendixes contain information about the test construction and validation process and comparative performances of various groups. A glossary is attached. (SLD)

RESULTS OF THE 1992 PRIMARY ASSESSMENT PROGRAM in urban schools

CURRICULUM & ASSESSMENT • RESEARCH & EVALUATION REPORT •

Report 3
1993

RESULTS OF THE 1992 PRIMARY ASSESSMENT PROGRAM in urban schools



Prepared for the Northern Territory Board of Studies
by the ERA Section of the Northern Territory Department of Education

Northern Territory Department of Education
Darwin 1993

IBSN 0 7245 2637 4

This book remains the property of

.....

For further information please contact:
Principal Research Officer
Evaluation, Research and Accreditation Section
Curriculum and Assessment
Northern Territory Department of Education
GPO Box 4821
Darwin NT 0801
Telephone (089) 89 5511

Printed for Publication Services, Northern Territory Department of Education by the Government Printer of the Northern Territory
P&P93/675–300

Acknowledgment:
Special thanks are due to members of the PAP team and members of various panels who developed the assessment materials.

Compilation: Gerry McCue

Results of the 1992 Primary Assessment Program in urban schools / prepared for the Northern Territory Board of Studies by the ERA Section of the Northern Territory Department of Education. Darwin : Northern Territory Department of Education, 1993.

ix, 53 p. : ill., forms ; 30 cm. (Curriculum & Assessment Research & Evaluation Report ; report 3, 1993.)
Compilation: Gerry McCue.
ISBN 0724526374 CIP

1. Education, Primary—Northern Territory—Evaluation. 2. Educational evaluation—Northern Territory. 3. Mathematical ability—Testing. 4. Reading (Primary)—Ability testing. 5. Urban schools—Northern Territory—Evaluation. I. McCue, Gerry. II. Northern Territory. Curriculum and Assessment Division. Evaluation, Research and Accreditation Section. III. Northern Territory. Board of Studies. IV. Series: Research and evaluation report (Northern Territory. Curriculum and Assessment Division) ; 3/1993.

372.126'099429 20

FOREWORD

For the first time in recent years the reports on the performances of urban and non-urban students on the annual PAP tests have been published as separate documents. This report describes the aggregated reading and mathematics performances of Years 5 and 7 students in urban Northern Territory schools.

Aggregated results for Years 5 and 7 students are encouraging, with apparent gains in some of the reading tests over the preceding year and no evidence of any declines in the others. Reading test results were especially pleasing; it would seem that the majority of NT Year 5 and 7 students were well able to read and comprehend texts considered suitable to their levels.

An aspect of this report that educators should find interesting is the detailed description of performance that is given in specific areas. Strengths and weakness have been highlighted. For example, in mathematics, both Year 5 and 7 students generally found activities involving perimeter difficult, but had demonstrated high levels of competence in basic algorithms and questions related to volume.

This year's report, for the first time, looks at test scores of boys and girls separately at Years 5 and 7. There is evidence that boys marginally out-performed girls in the measurement strand of mathematics. Results from the reading tests showed that in all the tests on reading, girls performed better than boys.

Test results of urban, self-identified, Aboriginal students indicated that, on average, their achievement levels were well below those of non-Aboriginal students. This gap was evident in all areas and the results indicate that in some areas the gap is larger at Year 7 than at Year 5. However, significant numbers of Aboriginal students did achieve at high levels in terms of raw test scores and in terms of scoring above the average scores of non-Aboriginal students.

I would take this opportunity, on behalf of the Northern Territory Board of Studies, to thank all those involved in the development, administration and reporting of the PAP urban tests for 1992. The program continues to grow in value each year and that is primarily owing to the assistance given in its design and implementation and to the good professional use schools are making of the results.

Dr CH Payne
Chairman
Northern Territory Board of Studies

CONTENTS

TABLES

GRAPHS

FIGURES

EXECUTIVE SUMMARY

Primary Assessment Program (PAP) tests for urban schools are conducted annually across the Northern Territory in reading and mathematics. This year's PAP tests indicated that, in terms of system performance on link questions, there was no difference between 1991 and 1992 performances in Year 7 reading comprehension and in Year 5 and Year 7 mathematics. Using the same analysis there were significant improvements in system performance on the Years 5 and 7 RDP tests and the Year 5 reading comprehension test.

Reading and mathematics tests developed as part of the PAP provide diagnostic information for teachers and schools and, as described above, allow performance at the system level to be monitored. Writing moderation, which facilitates comparability of teacher assessments of student writing in the NT, is the other aspect of the PAP.

The final chapter in this report is intended to foreshadow and describe the projected introduction of item response theory to contribute to the analysis and reporting of PAP results. Such changes will improve reporting methods and bring the NT into line with the larger Australian states. Results up till now have been analysed using traditional reporting methods. It is envisaged that next year's results will be analysed and reported using traditional methods and item response theory, thus enhancing the breadth and depth of analysis and at the same time maintaining the ease of understanding of results that is a feature of PAP reporting to its clients.

Aggregated system scores for the reading subtests were universally high. Within the subtests the most difficult questions were often of an inferential nature or questions where the answers were embedded in the text but required that some time and careful thought be devoted to their location.

It seems that generally students were able to read and comprehend the meanings of texts deemed by teachers to be appropriate to their level, i.e. Stage 5 for Year 5 and Stage 7 for Year 7. The same applies to the everyday texts that make up the reading for different purposes (RDP)) tests. Everyday texts include such items as dictionaries, phone books, newspapers, maps, timetables, product instructions and advertisements.

Results for the mathematics tests indicate that students found these somewhat more difficult than the reading tests. The average mark for both tests was around 55%.

At Year 5, students found measurement questions the most difficult and space questions the least difficult. Calculating perimeters proved difficult as did place value questions. As well as simple algorithm questions, students found questions related to capacity much less difficult.

Year 7 students found the three strands (number, measurement and space) to be of equal difficulty. Each of these strands had an average proportion correct of 55%. Questions with relatively simple mathematics but presented as 'problems' were found difficult as were perimeter and area questions. The relationships between units of mass also posed problems. At both Year 5 and Year 7, symmetry concepts, perimeter measurements and identifying and counting vertices, face and edges were not done well. Year 7 students found direction and time related questions to be of little difficulty.

Girls scored higher than boys in every reading subtest. Historically, this result was to be expected. With the Year 5 and Year 7 number and space strand questions there was very little difference between the genders but the boys' average scores for the measurement questions were clearly higher than those of girls for both tests. With all of the reading tests, greater proportions of girls than boys achieved scores in the higher ranges of the scoring continuum. This trend was reversed with the mathematics tests although the differences were not so large.

In terms of the PAP reading tests, the superiority of Aboriginal girls over their male counterparts was significantly larger than for the general population of Year 5 and Year 7 students. This superiority carried over to the mathematics tests where girls also out-performed boys but by lesser amounts.

On the cover of each test students were asked if they identified themselves as Aboriginal. Through this process around 10% of students sitting the PAP tests identified themselves as Aboriginal. On all tests the average for non-Aboriginal students was higher than for their Aboriginal peers although large numbers of Aboriginal students registered scores in excess of the average score for non-Aboriginal students. This trend was more evident in the reading test results. Distribution of scores in this report indicates that, once again especially with reading tests, there are high numbers of Aboriginal students achieving in the higher score ranges. In statistical terms, there was no improvement over the Aboriginal students who had sat the equivalent PAP tests the previous year, and in a number of cases performance dropped. See Appendix II for further details.

With the Year 5 mathematics test the average scores of Aboriginal students were 20% below those of non-Aboriginal students. For Year 7 this gap was 27%. This trend applied also to the reading comprehension tests but with the RDP tests the average performance of Aboriginal students, whilst still behind, was closer to non-Aboriginal students at Year 7 than at Year 5.

INTRODUCTION

Aims of the Primary Assessment Program (PAP)

The PAP's aims in urban schools are to annually monitor system-wide performance in reading and mathematics and to report student, school and system results. 'System' refers to all of the schools and therefore all students participating in the testing program.

The PAP's aims are as follows

at the classroom level

- to provide test results of individual students in reading and mathematics
- to give teachers bases for improving their teaching programs and assisting individual students

at the school level

- to provide test data that would enable schools to compare both individual and school results with aggregated Territory (system) results
- to give schools data on which they can base program improvements
- to provide schools, over a period of time, with a bank of high quality and appropriate assessment materials

at the system level

- to detect trends in student performance over the years
- to report on the performance of Territory students in mathematics and reading
- to identify areas of strength and weakness in mathematics and reading
- to inform planning curriculum development.

Background to the Tests

PAP tests are conducted by the Northern Territory Department of Education's Curriculum and Assessment Division under the auspices of the Northern Territory Board of Studies. This system-wide testing program for Years 5 and 7 in urban schools commenced in 1983. The testing program, using different tests, was widened to include non-urban (Aboriginal) schools in 1986.

Policy matters are the province of the Primary Assessment Committee (PAC) which monitors this constantly evolving program and advises the NT Board of Studies. The PAC responds to suggestions from teachers, principals, school councils, administrators and other interested parties. In large part, the PAP's recognition and

acceptance is due to its willingness to make changes in response to the needs of its client groups.

Not to be overlooked are the great numbers of teachers who, under the direction of the experienced PAP team, work hard at inservices and test construction panels developing quality test materials.

PAP Assessment

The four components of the PAP are

- mathematics tests
- reading comprehension tests
- the reading for different purposes (RDP) tests
- writing moderation in a range of genres

All Year 5 and 7 students are tested annually in reading and mathematics. Schools have one week to complete the tests. Time to complete each test is recommended rather than prescribed. The moderation process is an ongoing exercise.

Writing Moderation

This year schools were required to send samples of students' writing in the report and narrative genres, which had been assessed according to the assessment frameworks provided. These frameworks show students' attainment in writing at Stages pre-3, 3, 5, 7 and beyond 7.

The assessments were then moderated by panels of experienced school and office-based teachers who also selected a range of student writing for inclusion in the moderation booklets. These booklets are used for assessment and moderation at the school and system levels.

Test Construction

In having tests written by panels comprising school and office-based teachers, questions can be pitched as close as possible to the appropriate levels of difficulty. Instructional validity is maintained through careful matching of instructional objectives with test objectives. Extensive trials are then held of all potential questions and inappropriate or ambiguous questions can then be modified or eliminated. All tests comprise multiple choice questions where students are given four options to choose from and open-ended questions where the students are asked to supply a short answer.

The mathematics questions are constructed to conform to a chart of specifications that has been derived from the Western Australian *Learning Mathematics* syllabus, which is used in all NT schools. This ensures that the spread of test questions is commensurate with the syllabus, for example approximately half of the syllabus is devoted to the number strand so approximately half the test questions come from this strand. A specified number of these questions are designed to be 'easy', 'medium' or 'hard'. Finally, questions are designed to test levels of understanding of mathematical concepts ranging from recall, computation, understanding to application, with an emphisis on understanding and application.

Trialling of Tests

All tests are extensively trialled in a range of schools prior to finalisation. From trial results it is possible to estimate the level of difficulty, reliability and the discrimination factor for each question. This information is presented to a test review panel and questions are then modified or eliminated where necessary.

Link Questions

A statistical analysis is carried out to compare system performance from one year to the next. This is achieved by the use of 'link' questions, i.e. common questions used over succeeding years. It should be mentioned that as only a small number of questions are used as links, care should be taken in interpreting changes in performance from one year to the next, although these questions do provide a good guide.

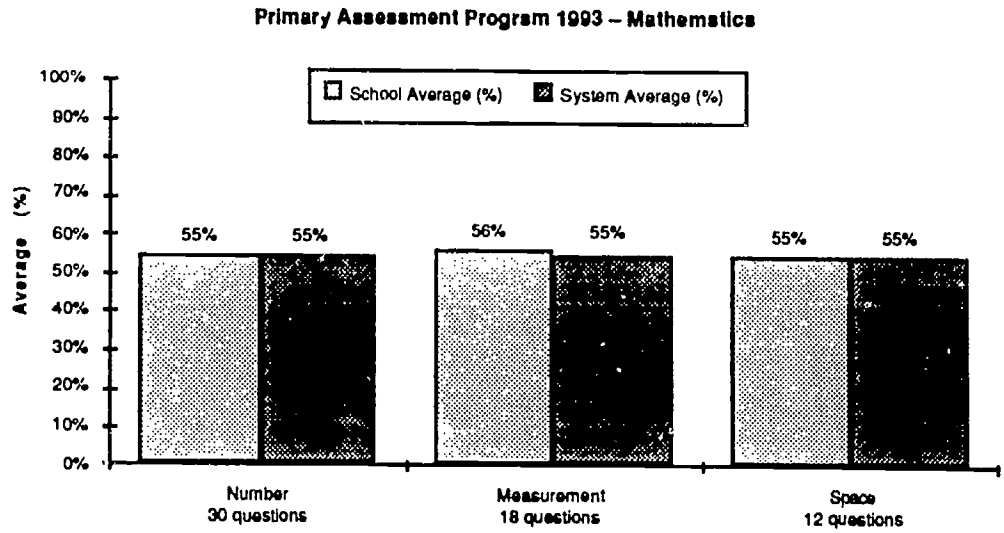
Reporting Results

All individual and school results are confidential and any public reporting is of aggregated system results only. Each school receives reports for each of the tests done by its students and may use these results in ways they feel are appropriate. School councils also receive summary statistics of their school's results.

Result booklets containing individual, school and system results are sent to schools. To enhance the diagnostic capacities of the testing program, results are set out in such a form that teachers are able to easily read and interpret the graphs and tables of statistics. A section of this booklet deals with how to read and interpret the statistics.

Graphs

Graphs provided to each school show the average subtest scores for that school compared to the system averages for the subtests. Each test is made up of a number of subtests. In mathematics there are three subtests, which are the number, space and measurement strands. In the reading tests, texts and accompanying questions are regarded as subtests.



Tables

School Report 1

This series of tables shows, for each question, the school's performance compared to the system. Also included is a concise description of the task required by each question. School Report 1 also indicates whether the difference between the school's and system performance on each question was statistically significant or not.

Primary Assessment Program– 1992 Test Results
School Report No. 1: Comparative Data

School ID
SAMPLE

Testcode : urec592

	SCHOOL	SYSTEM	DIFFERENCE
31 MAKE INFERENCE ON THE BASIS OF INFORMATION READ	.86	.81	.06
32 DETERMINE WORD MEANING FROM CONTEXTUAL CLUES	.95	.91	.04
33 IDENTIFY THE PURPOSE OF A REPORT	.88	.81	.07
34 IDENTIFY THE MAIN IDEA OF A PARAGRAPH	.90	.86	.04
35 IDENTIFY THE MAIN IDEA OF A PARAGRAPH	.95	.90	.05
36 IDENTIFY THE MAIN IDEA OF A PARAGRAPH	.95	.87	.08
37 MAKE INFERENCE ON THE BASIS OF INFORMATION READ	.53	.51	.02
38 LOCATE INFORMATION	.80	.76	.04
39 UNDERSTANDINGS THE GENERIC 'THE'	.85	.76	.09
40 LOCATE INFORMATION	.93	.83	.10
Statistics			
Number of Items	15.00	15.00	
Subtest Mean	12.75	11.86	.89
Standard Deviation	2.96	3.15	
t-value (significant)	2.27	.00	
TOTAL TEST			
Statistics			
Number of Items	40.00	40.00	
Number of Students	59.00	1975.00	
Total Mean	31.86	30.97	.90
Standard Deviation	7.17	7.18	
t-value (not-significant)	.95	.00	

School Report 2

For each of the three tests this report gives individual results and system rankings for each student expressed as percentile ranks.

Primary Assessment Program-1992 Test Results
School Report 2: Individual Results and System Ranking

SCHOOL CODE: SAMPLE
TESTCODE: UREC592
NO. OF ITEMS: 40
NO. OF STUDENTS: 1975

STUDENT NUMBER	NUMBER CORRECT	% CORRECT	PERCENTILE RANK IN NT
7	37	92.50	86.025
50	37	92.50	86.025
55	37	92.50	86.025
501	37	92.50	86.025
10	38	95.00	92.861
37	38	95.00	92.861
45	38	95.00	92.861
60	38	95.00	92.861
32	39	97.50	98.127
47	39	97.50	98.127
51	39	97.50	98.127
56	39	97.50	98.127
58	39	97.50	98.127
4	40	100.00	100.00
59	40	100.00	100.00
42			ABSENT

Test Reliability

Estimates of test reliability were derived using the Kuder-Richardson Formula 20 (KR 20) calculation. See Table 1.

Table 1 Reliability Estimates of PAP Tests Using the KR 20 Formula

YEAR	Reading Comprehension	RDP	Mathematics
5	0.90	0.93	0.92
7	0.90	0.91	0.94

All the tests yielded high reliability estimates which means that one could expect student's scores to be consistent from one measurement to the next in each of the areas in which testing was conducted. This has been achieved through carefully assembling questions with known characteristics obtained through the trials.

Mathematics Tests

Students at Years 5 and 7 are given a sixty-question mathematics test based on the Western Australian *Learning Mathematics* syllabus. Each test has two thirty-question sections. Test design specifications ensure an appropriate balance of number, space and measurement questions as well as a range of difficulty and a mixture of recall, computation, understanding and application questions. There is a number of link questions in each of these tests to provide a basis for comparing performance from one year to the next. See Figures 8 and 11 for sample questions.

Reading Tests

Two reading tests are set for Year 5 and 7 urban students, namely, a forty-question reading comprehension and a fifty-question RDP test. Questions are multiple-choice or open-ended. From 1993 the RDP test will be known as the reading everyday texts (RET) test. Sample RDP and reading comprehension texts and associated questions appear in Appendixes IV and V.

Both the Year 5 and 7 reading comprehension tests were forty questions which were derived, in both cases, from three texts. The fifty-question Year 5 and 7 RDP tests were based on eight texts at Year 5 and eleven texts at Year 7. In the comprehension tests students attempted a range of tasks such as locating information, drawing a conclusion on the basis of information read, identifying sequences of events etc. Broadly speaking the questions can be classified as interpretive or inferential although it must be said that some questions are difficult to so classify.

The RDP texts are presented as a newspaper. These are known as environmental texts and include materials that students may find in their everyday environments which impart information of a more functional nature. Such texts include cultural event posters, dictionary pages, cinema guides, public notices and safety warnings.

2. YEAR 5 RESULTS

Reading Comprehension Test Results

Table 2 **Overall Statistics for the Year 5 Reading Comprehension Test**

Number of Questions	40
Average Score	31
Number of Students	1975
Standard Deviation	7.18

This, as in 1991, was a forty-question test. More students sat the test in 1992 than in 1991 when 1892 students did the test. The average score in 1992 was 77% compared to 67% in 1991. Average score is calculated by summing individual scores then dividing this figure by the number of students. The resultant figure, which is a score out of forty, may then be converted to a percentage.

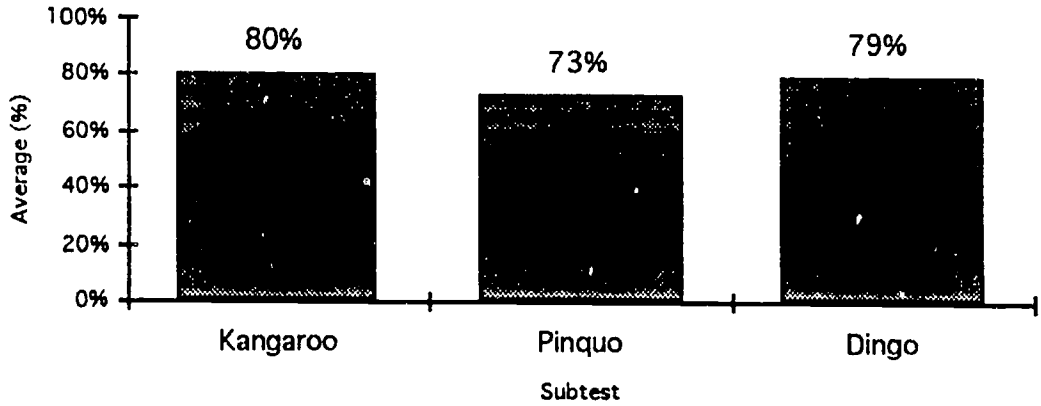
In this instance, the standard deviation's size indicates that individual scores varied by a reasonable amount, as was the case in the preceding year. Standard deviation quantifies how much scores varied from the average score. It is a measure of the spread of scores over the continuum of possible scores.

Thirty-seven students attained the maximum score of forty and eight students got zero.

Table 3 **Passages in the Urban Year 5 Reading Comprehension Test**

Subtest/Title of Passage	Genre
Kangaroo	Narrative
Pinquo	Descriptive
Dingo	Report

Graph 1 **Average Subtest Scores for the Year 5 Reading Comprehension Test**

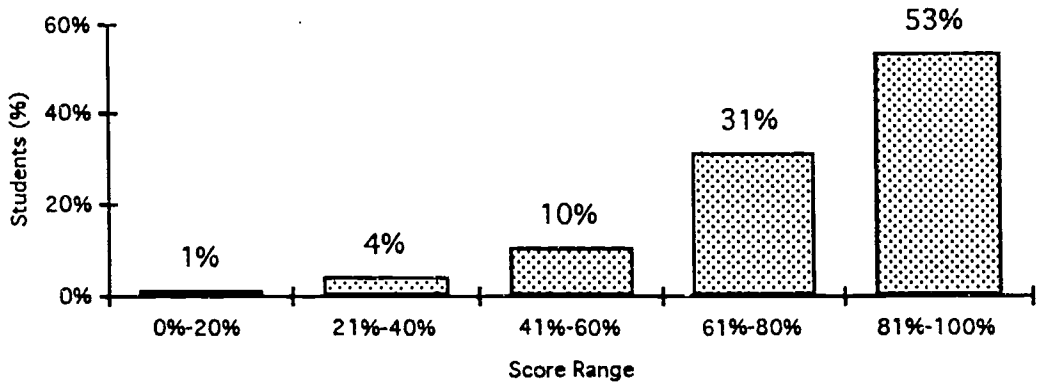


On the whole, Year 5 students demonstrated a high degree of comprehension of the three passages. A descriptive passage about a penguin, Pinquo, was the most difficult. Kangaroo, a Dreamtime narrative, and Dingo, which was a report, were easier.

Students found a question that involved placing four events in the sequence in which they occurred in the text the most difficult. Forty-six per cent of students were able to do so. Of the six most difficult questions, three were of an inferential nature and three were interpretive. Inferential questions made up about 20% of all questions which indicates inferential questions were over-represented at the difficult end of the scale.

Students achieved a score of 90% or better on seven questions.

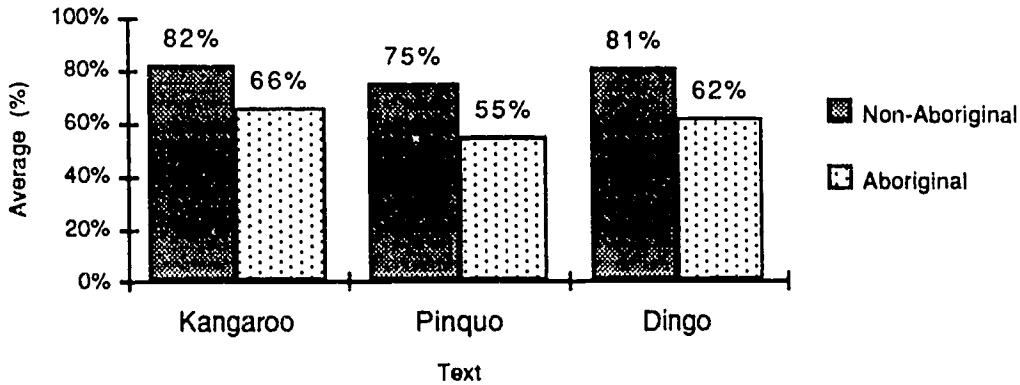
Graph 2 Distribution of Scores for the Year 5 Reading Comprehension Test



A distribution of scores graph describes what percentage of students who took a test achieved a score in each of the score ranges described. For example, in the above graph, 52.51%, or 1037, of the 1975 students achieved a score of 81% – 100% inclusive and 1.16% (23) scored less than 21.

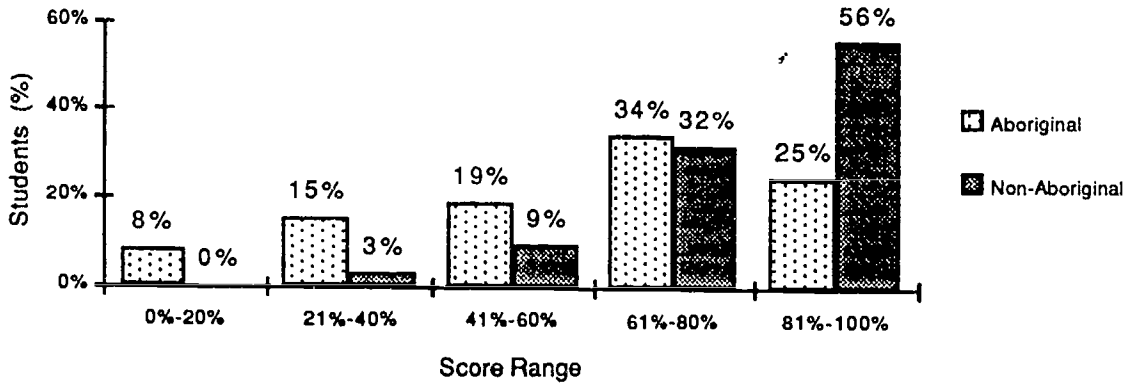
Eighty-four per cent of students achieved a score in excess of 61%. Scores were generally distributed towards the high end of the scale which could point to a generally good level of reading ability.

Graph 3 Average Scores of Non-Aboriginal and Aboriginal Students for the Year 5 Reading Comprehension Test



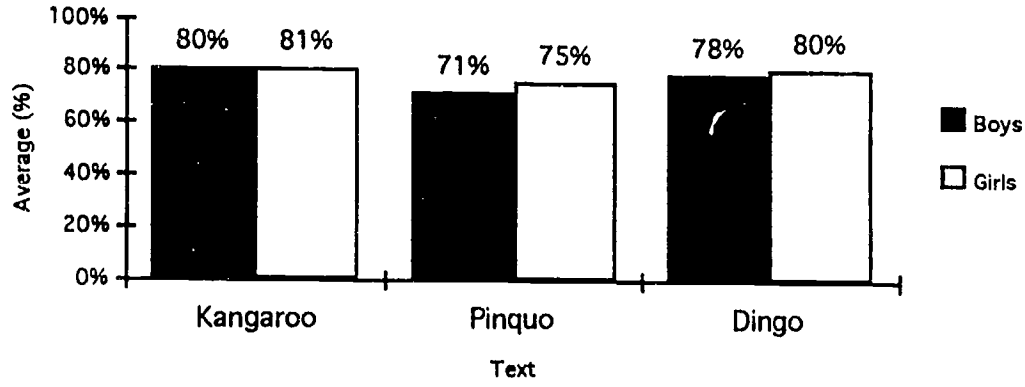
Graph 3 shows the average scores achieved by non-Aboriginal and Aboriginal students for each subtest. For example, with the Dingo subtest non-Aboriginal students averaged 81% and Aboriginal students 62%. Students were asked to indicate on the cover of the test papers if they were Aboriginal. About 10% of students sitting the PAP tests identified themselves as Aboriginal. Non-Aboriginal students performed better on this test than their Aboriginal counterparts. The average score for non-Aboriginal students was 79% and 61% for Aboriginal students. The biggest difference was seen in the most difficult subtest, Pinquo. Care should be taken when looking at such differences purely in terms of differences between average scores, as the following graph demonstrates.

Graph 4 **Distribution of Scores of Aboriginal and Non-Aboriginal Students on the Year 5 Reading Comprehension Test**



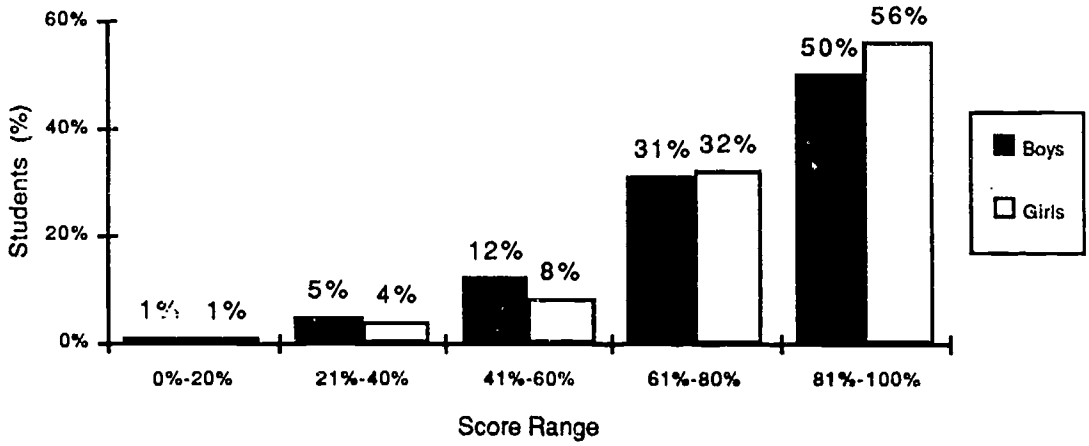
Fifty-nine per cent of Aboriginal students scored in excess of 60%. On average Aboriginal students did not perform as well as their non-Aboriginal peers and this is reflected in the above graph, although 32% of Aboriginal students achieved a score that was above the average score of non-Aboriginal students.

Graph 5 **Average Scores for the Year 5 Reading Comprehension Test by Gender**



Girls out-performed boys in all subtests with the biggest difference being in Pinquo, the most difficult subtest. On average the performance of girls exceeded that of boys by 3%.

Graph 6 **Distribution of Boys' and Girls' Scores on the Year 5 Reading Comprehension Test**



Higher proportions of girls than boys occupied the higher ranges of the scoring continuum. This difference was more exaggerated when the scores of Aboriginal girls and boys were analysed. Appendix IV contains further information on the distribution of scores expressed in terms of gender.

Scores on the Year 5 reading comprehension test were higher in 1992 than 1991. A t-test analysis revealed that the difference was statistically significant.

A significance level of $p<.05$ is used with all t-tests contained in this report.

See Appendix II for a further description of performance on link questions.

Reading for Different Purposes Test Results

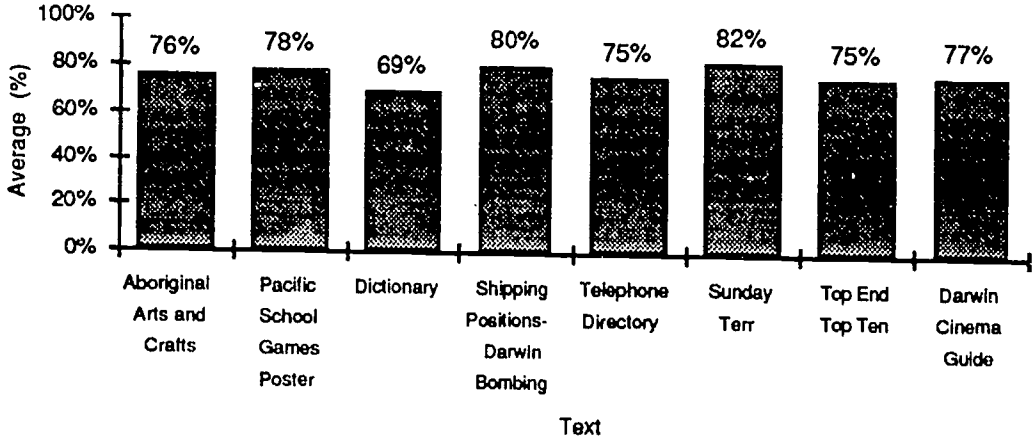
Table 4 **Overall Statistics for the Year 5 RDP Test**

Number of Questions	50
Average Score	38
Number of Students	1961
Standard Deviation	9.78

More students did this test than in 1991 when 1877 students did the test. The high standard deviation indicates that there was considerable variation in individual scores.

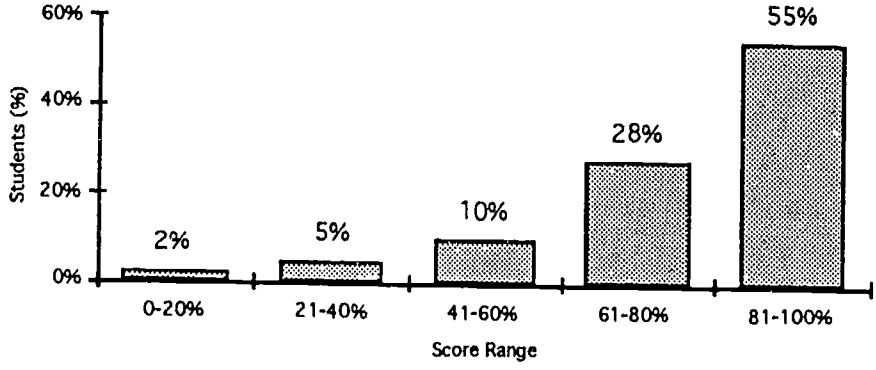
Seventeen students achieved the highest possible score and nine students failed to score.

Graph 7 **Average Subtest Scores for the Year 5 RDP Test**



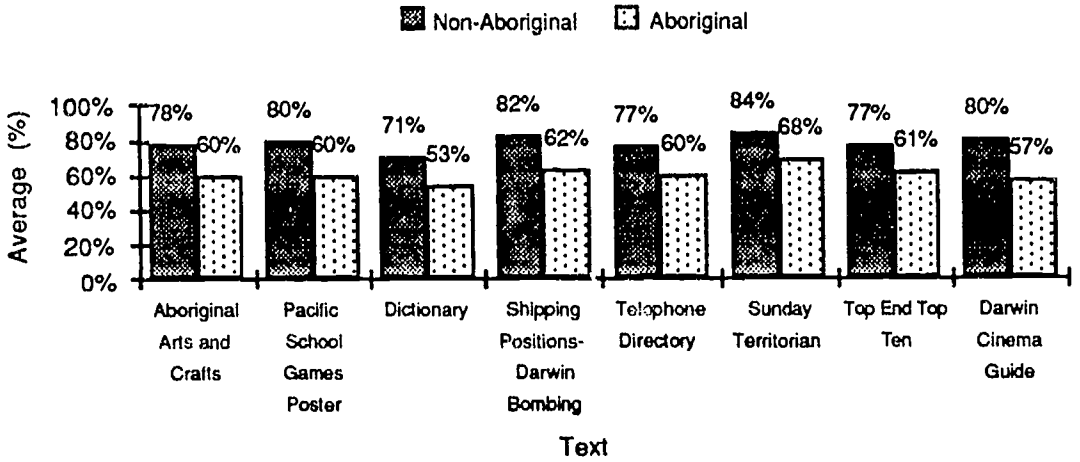
Many of the questions which the students found the most difficult were those which, although not actually difficult in themselves, involved searching through a large amount of text. In addition, some of the multiple choice options were very close in meaning and therefore required time and careful consideration in selecting the correct one. The most difficult question involved students first having to know the meaning of a term ('guide word') in order to answer the question.

Graph 8 **Distribution of Scores for the Year 5 RDP Test**



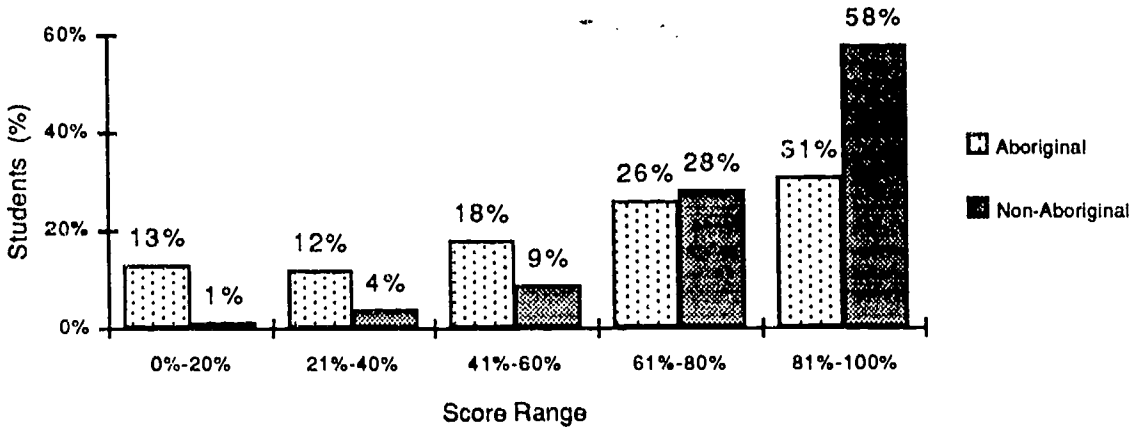
Of the 1961 students who sat this test, 55.1% (1081) scored in excess of 81% correct. Eighty-three per cent of students scored over 60% which seems to indicate most students were able to easily understand and answer questions relating to the everyday texts such as a dictionary excerpt, advertisements and articles from newspapers, and telephone directory information.

Graph 9 **Average Scores of Non-Aboriginal and Aboriginal Students for the Year 5 RDP Test**



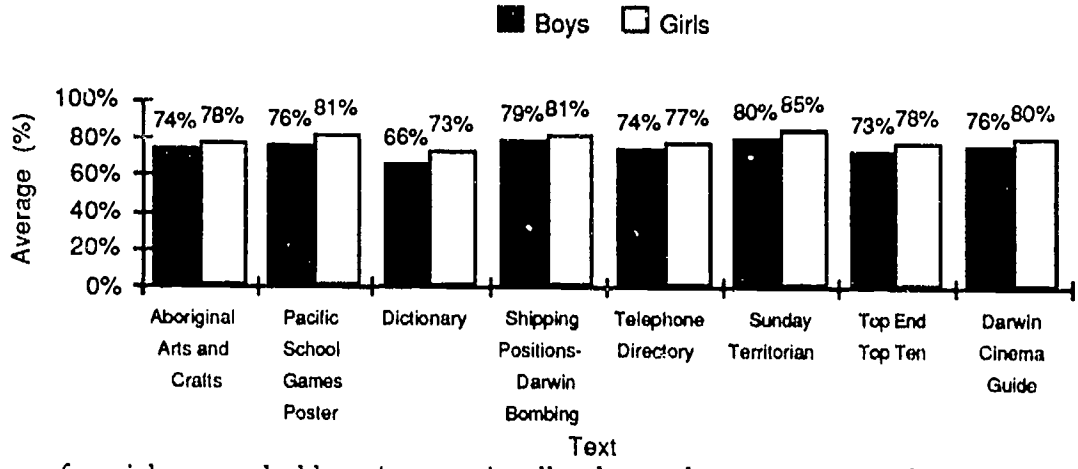
Average scores of non-Aboriginal students exceeded those of Aboriginal students in all subtests by an average of 18.5%. Within the subtests, average differences ranged from 16% for both the *Sunday Territorian* page and the account of the Top End Top Ten, to 23% for the Darwin Cinema Guide. The difference in scores between the two groups did not seem related to the difficulty of the subtest.

Graph 10 **Distribution of Scores of Aboriginal and Non-Aboriginal Students on the Year 5 RDP Test**



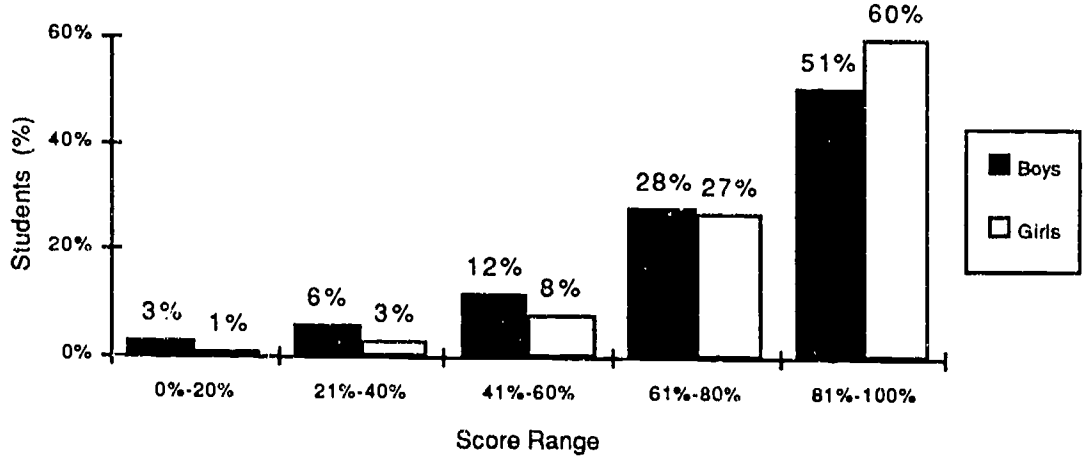
From the above graph it may be seen that the majority of Aboriginal students achieved scores in the higher ranges of the scoring continuum. System results show that 32% of Aboriginal students who sat this test scored higher than the average for non-Aboriginal students.

Graph 11 **Average Scores for the Year 5 RDP Test by Gender**



Scores for girls exceeded boys' scores in all subtests by an average of 4.4%. These differences ranged from 2% for questions based on the Shipping Positions During the First Bombing of Darwin map and information sheet, to 7% for questions based on a dictionary excerpt.

Graph 12 **Distribution of Boys' and Girls' Scores on the Year 5 RDP Test**



An analysis of the distribution of boys and girls scores revealed that 79% of boys scored 61% or higher while 87% of girls scored in this range. With Aboriginal students 12% more girls than boys scored in this range. See Appendix IV for a more detailed look at these aspects of test performance.

Scores on the link items in the Year 5 RDP test were higher in 1992 than 1991. A t-test analysis revealed that the difference was statistically significant.

See Appendix II for a further description of performance on link questions.

Mathematics Test Results

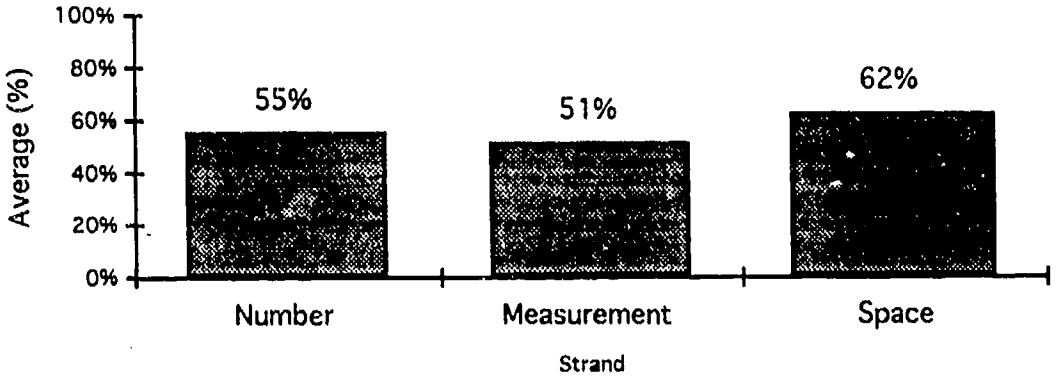
Table 5 Overall Statistics for the Year 5 Mathematics Test

Number of Questions	60
Average Score	33
Number of Students	1971
Standard Deviation	11.84

A quite large standard deviation for this test points to a relatively large variation in individual scores. Marginally more students sat this test than in 1991 when 1924 students did the test.

No students achieved the maximum possible score of sixty. Three students achieved the highest score of fifty-nine out of sixty. There were no questions with zero per cent correct. Six students scored zero.

Graph 13 Average Subtest Scores for the Year 5 Mathematics Test



Students clearly found questions from the space strand easiest, followed by number, then measurement questions where the average score was 51%. In last year's mathematics test students found number questions easiest (61%) followed by space (58%), then measurement (53%). Such fluctuations over successive years can be caused by the inclusion of some questions that turn out to be very easy or very difficult.

Figure 1 Year 5 Number Strand Level of Difficulty Analysis

SCORE RANGE*	QUESTION NUMBER	DESCRIPTION OF TASK
0–25%	6	Demonstrate part/whole nature of fractions
	24	Write number to two decimal places in alternate form
26–50%	44	List all factors for 18
	49	Multiply decimals (cents) by whole number: .55 x 48
	18	Construct and use codes
	22	Add/subtract length in metres to three decimal places
	21	Show relationships between places in base 10
	58	Problem solving: combinations of 2 from a group of 4
	40	Divide money by whole number to 10: $288 ÷ 8
	9	Investigate factors of 24
	3	Multiply number <100 with regrouping: 39 x 15
	50	Problem-solving with multiples and Venn diagram
	15	Subtract a fraction from a whole number: 1 - 3/5
	42	Give most frequent score (mode)
51–75%	32	Multiply numbers: 836 x 7
	34	Show understanding of place value
	47	Insert sign to complete number sentence: 3.1 > 3.01
	54	Complete a number sentence involving fractions
	20	Add numbers to three decimal places
	1	Subtract number to three decimal places: 3.850 kg - 2.435 kg
	14	Identify largest group from standard tally
	2	Add whole numbers: 216 + 97 + 8 + 504
	37	Identify a square number (36)
	13	Interpret data using standard tally
	17	Order three numbers that have three decimal places
	29	Interpret data represented by a bar graph
	45	Solve a problem using multiplication: $12 x 9
	41	Find middle (median) score
76–100%	31	Add whole numbers: 659 + 78
	10	Solve a problem using multiplication: $4.50 x 3

* Score Range shows the per cent correct category into which each question falls. For example, between 0 – 25% of students got Questions 6 and 24 correct. These were therefore difficult questions.

Of the thirty questions in the number strand, fourteen (47%) were in the 0–50% score range. For two questions, 76% or more students gave the correct answer. Questions involving decimals made up a number of the more difficult questions but questions involving decimals were also well represented among the easier questions . The two factor questions in the test proved difficult. It was interesting that what appeared to be a reasonably simple algorithm (39 x 15) in question 3 proved quite difficult. In contrast the easiest number strand question was also a relatively simple multiplication that was presented in problem form and involved decimals, i.e. dollars and cents.

Figure 2 Year 5 Measurement Strand Level of Difficulty Analysis

SCORE RANGE	QUESTION NUMBER	DESCRIPTION OF TASK
0–25%	59	Calculate perimeter of a polygon
	11	Determine mass in kg of objects on a scale
26–50%	56	Calculate perimeter of shape drawn on a grid
	52	Add masses of two objects in g and express in kg: 700 g + 900 g
	55	Compare perimeters of shapes drawn on a grid
	12	From airline schedule calculate total duration of stops
	51	Transfer digital clock time to clock face
	7	Determine volume by counting cubes in a shape
51–75%	30	Solve problem involving the comparison of volumes
	43	Draw 2D shape, area specified, on grid
	16	Find area of irregular shape by counting squares
	27	Convert time interval, expressed in min/sec, to sec
	39	Convert m (with two decimal places) to cm: 2.04 m
	23	Recall capacities (mL, L) of common containers
	48	Compare 3D objects to determine largest volume
	33	Relate length/volume measurement to mass
76–100%	5	Mark specified volume (mL) on a measuring jug
	36	Compare capacities

Eight (44%) of the measurement strand questions resulted in 50% or less students giving a correct answer and two questions (11%) being in the score range of 76% or better. Three of the five hardest questions involved the calculation of perimeters. The other two were mass related questions. The five easiest questions involved concepts related to capacity.

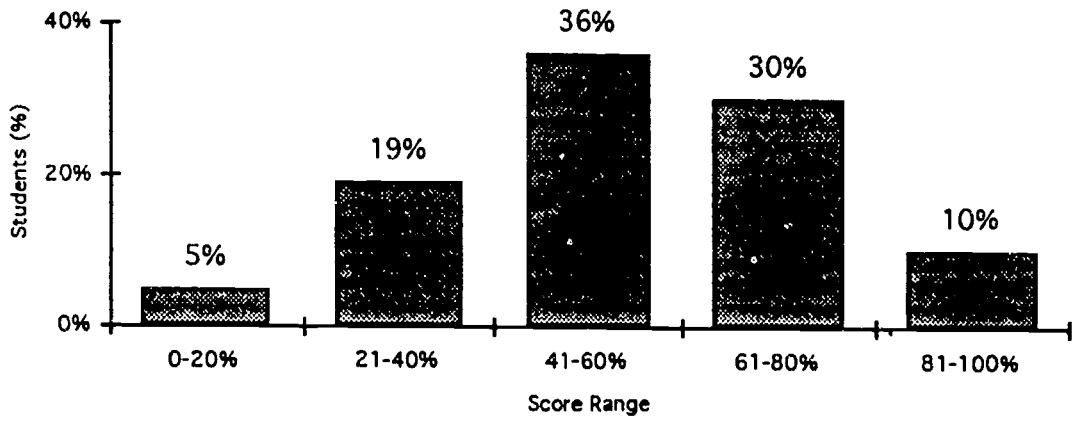
Figure 3 Year 5 Space Strand Level of Difficulty Analysis

SCORE RANGE	QUESTION NUMBER	DESCRIPTION OF TASK
0–25%	28	Count vertices on a rectangular prism
26–50%	46	Halve the size of a shape on a grid
	19	Draw all lines of symmetry for given shape
51–75%	53	Identify the right-angled circle sector
	38	Locate coordinates to spell a word
	57	Investigate changes of position to complete a pattern
	35	Enlarge by factor 2 a shape on a grid
	4	Recognise the tessellated shape
	26	Complete a grid diagram to make shape symmetrical
	25	Determine position after following directions

76–100%	8	Investigate shapes and symmetry
	60	Investigate layout of environment

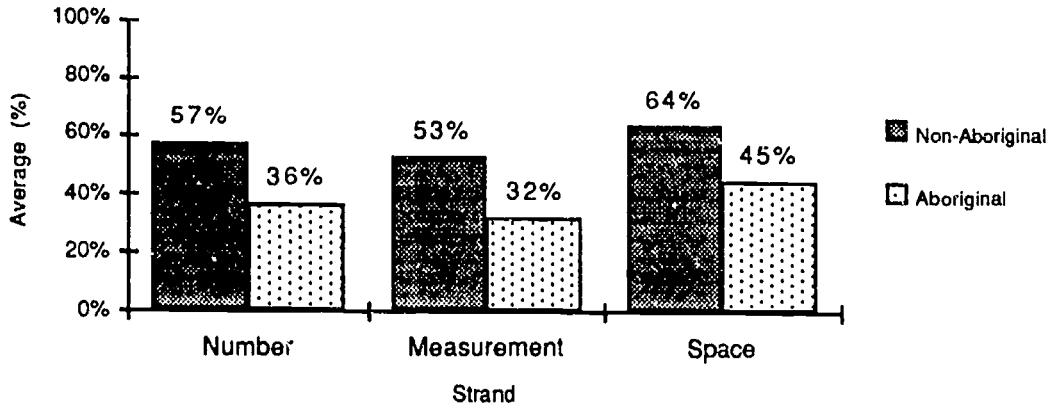
Three (25%) of the space questions had a per cent correct of 25% or less while two (17%) rated 76% or better. Two of the three lowest scoring questions required students to understand the terms vertices and symmetry. The second easiest question partly involved the concept of symmetry but the term symmetry was not used. The easiest question involved determining how many triangles made up a hexagon.

Graph 14 Distribution of Scores for the Year 5 Mathematics Test



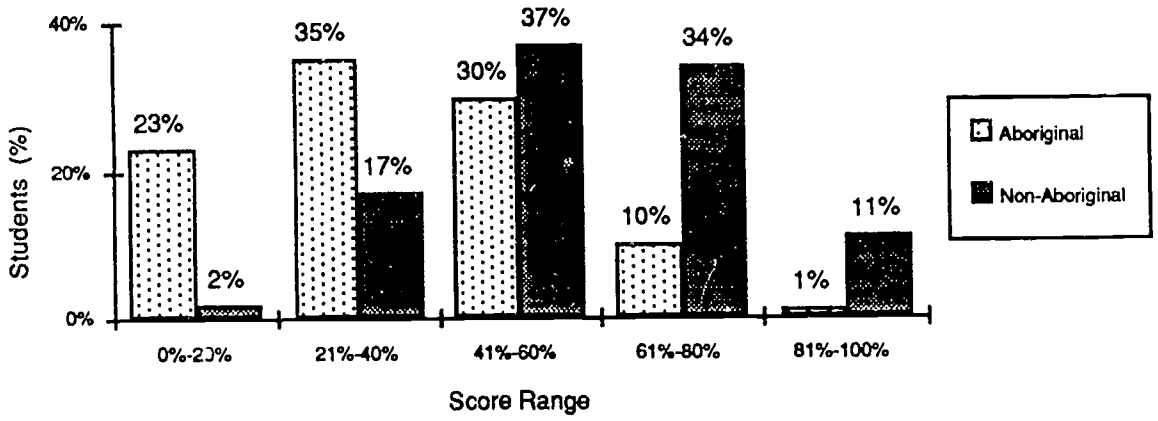
Unlike the distributions of the two reading tests, the distribution of scores for the mathematics test closely resembles the shape of the normal curve. The normal curve occurs when most scores are clustered around the average score and taper towards the low and high ends of the score range continuum. In this case 714 students were in the 41% – 60% range, ninety in the 0 – 20% range and 196 were in the 81% – 100% range.

Graph 15 Average Scores of Non-Aboriginal and Aboriginal Students for the Year 5 Mathematics Test



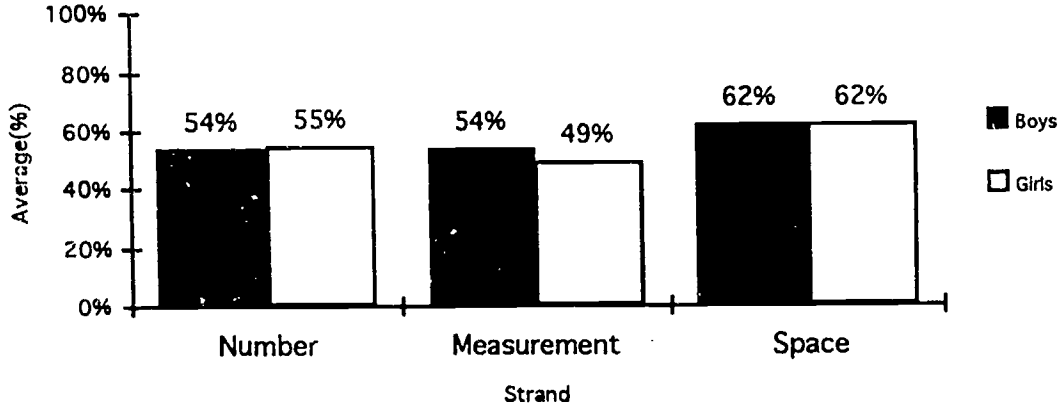
In terms of the overall test, the average scores for Aboriginal students and non-Aboriginal students were 37% and 57% respectively. Strand-wise, the biggest difference (21%) between these groups occurred in both number and space.

Graph 16 **Distribution of Scores of Aboriginal and Non-Aboriginal Students on the Year 5 Mathematics Test**



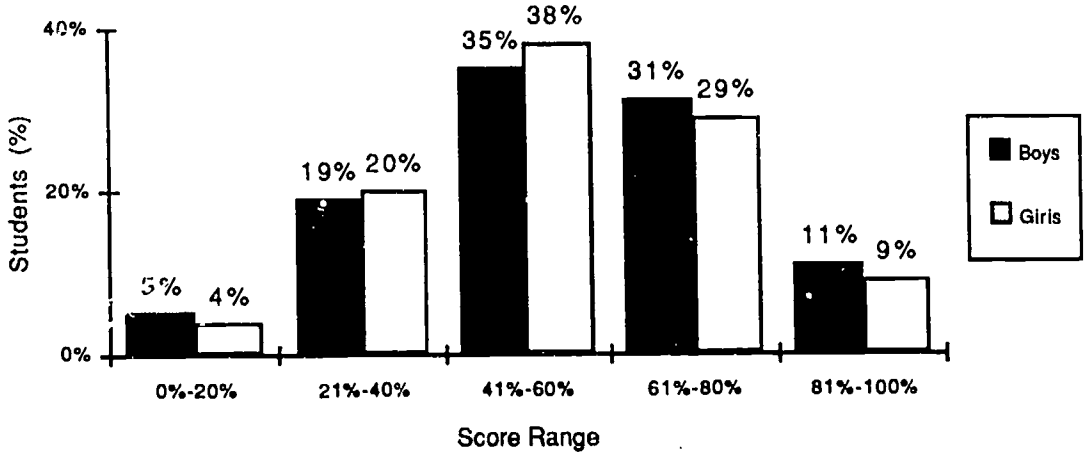
In the Year 5 mathematics test most Aboriginal students were placed towards the lower end of the score range scale. Fifteen per cent of Aboriginal students achieved a score that exceeded the average score of non-Aboriginal students.

Graph 17 **Average Scores for the Year 5 Mathematics Test by Gender**



Girls, who did slightly better than boys in number questions, were out-performed by boys by 5% in measurement questions. Boys and girls scored the same for space questions.

Graph 18 **Distribution of Boys' and Girls' Scores on the Year 5 Mathematics Test**



Unlike with the two reading tests, the mathematics test saw a higer proportion of boys in the higher scoring ranges than girls; although this was not the case with Aboriginal students where 8% of boys and 16% of girls scored in excess of 60%. See Appendix IV more detailed information on the distribution of scores looked at in terms of gender.

Analysis of Year 5 mathematics questions using a t-test showed that the difference between 1992 and 1991 performances on the link questions was not statistically significant.

See Appendix II for a further description of performance on link questions.

3. YEAR 7 TESTS

Reading Comprehension Test Results

Table 6 Overall Statistics for the Year 7 Reading Comprehension Test

Number of Questions	40
Average Score	29
Number of Students	1782
Standard Deviation	7.78

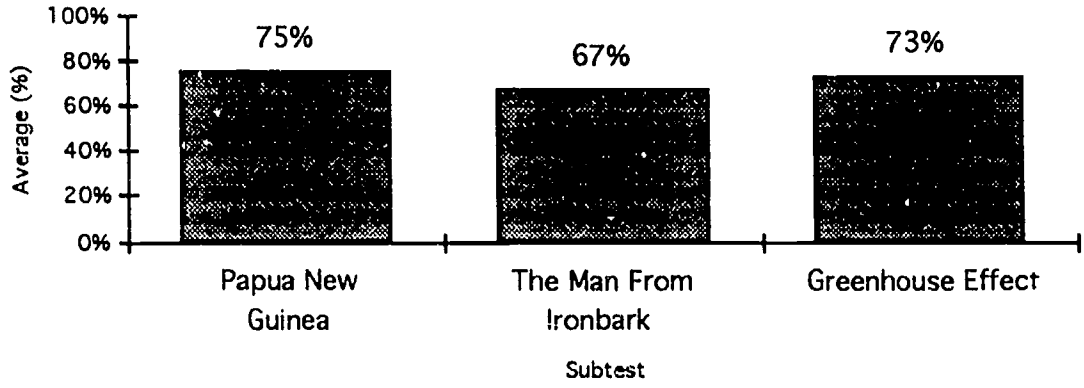
In 1991 1842 students sat the reading comprehension test. Once again the reasonably high standard deviation indicates some variation around the average score. An overall average of 72% shows that most students were able to achieve well on this test.

Thirty-two students scored forty. Twenty-three of these were girls. Four students scored zero.

Table 7 Passages in the Urban Year 7 Reading Comprehension Test

Subtest/Title of Passage	Genre
Papua New Guinea	Report
The Man From Ironbark	Ballad
Greenhouse Effect	Explanation

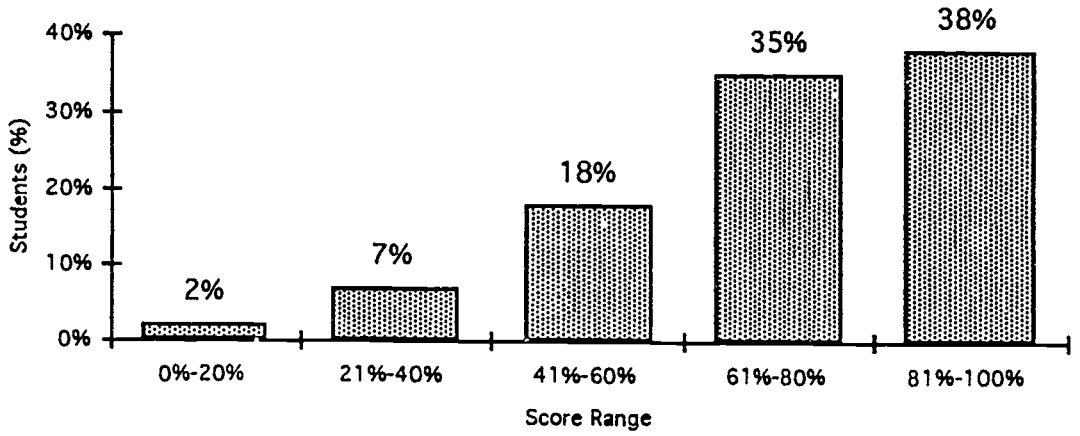
Graph 19 Average Subtest Scores for the Year 7 Reading Comprehension Test



In a test that featured reasonably high scores, students found the AB 'Banjo' Patterson ballad, *The Man From Ironbark* the most difficult subtest. The extensive use of colloquial language in this text may have contributed to this.

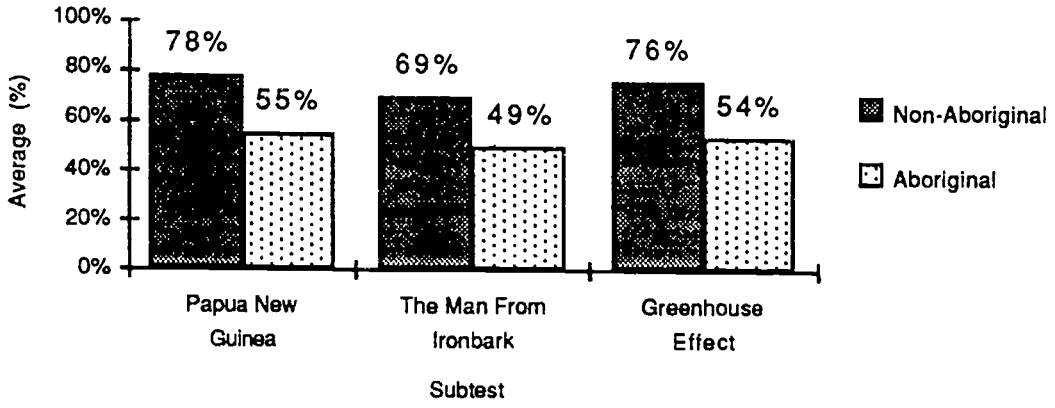
No obvious patterns emerged between the type of question asked and the level of difficulty as expressed by the per cent correct. It would appear that the level of difficulty of the text was the main determinant of level of difficulty of the question and not the question type. From the subtest average scores it can be seen that there was not a great deal of variation in levels of text difficulty.

Graph 20 **Distribution of Scores for the Year 7 Reading Comprehension Test**



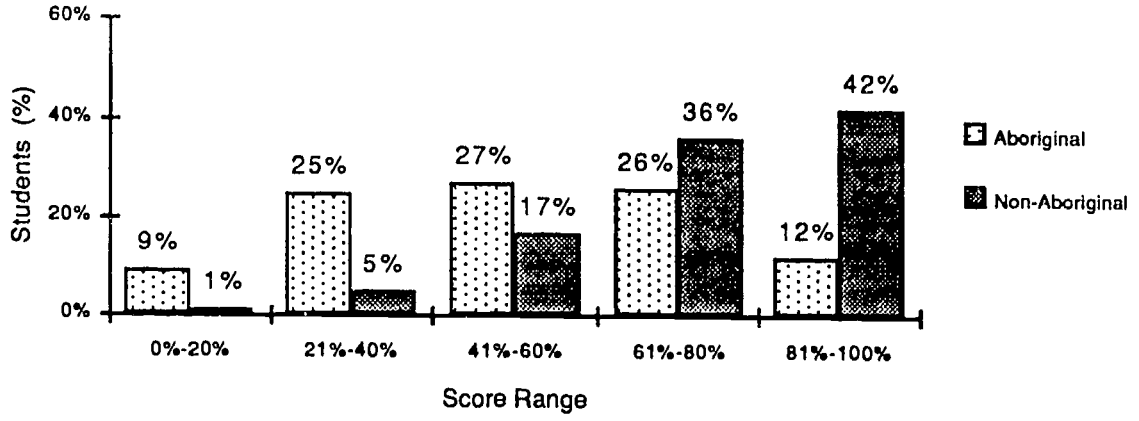
The skew to the right in the above graph indicates that most students who sat the test performed rather well, in fact 1301 of the 1782 students who took the test scored in excess of 61%.

Graph 21 **Average Scores of Non-Aboriginal and Aboriginal Students for the Year 7 Reading Comprehension Test**



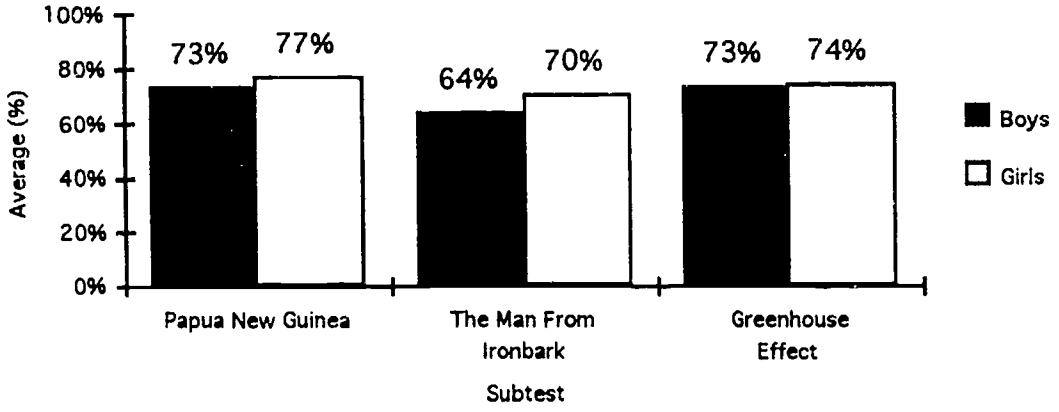
Non-Aboriginal students found all of the subtests easier than their Aboriginal peers. The largest difference of 23% occurred in the Papua New Guinea subtest which was a report on some of that nation's major geographical, economic and cultural characteristics. The average difference in scores was 21.7%.

Graph 22 **Distribution of Scores of Aboriginal and Non-Aboriginal Students on the Year 7 Reading Comprehension Test**



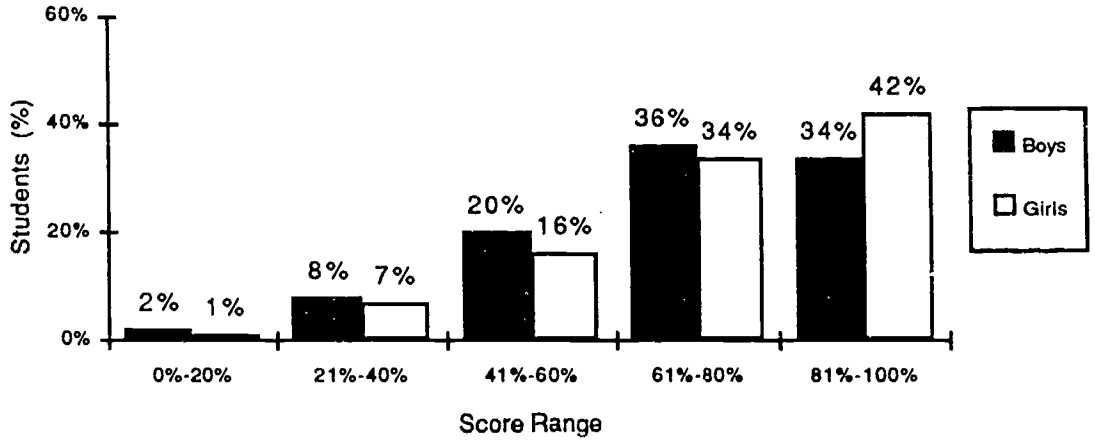
Graph 22 shows that 38% of Aboriginal students recorded a score of 61% or higher. Twenty-one per cent of Aboriginal students' scores exceeded the average score for non-Aboriginal students.

Graph 23 Average Scores for the Year 7 Reading Comprehension Test by Gender



Girls performed better than boys in all subtests. When an explanation text of a more scientific nature was used, i.e. Greenhouse Effect, the performance of boys came closest to that of girls.

Graph 24 Distribution of Boys' and Girls' Scores on the Year 7 Reading Comprehension Test



With system and also Aboriginal students' scores, 6% more girls than boys achieved a score of 61% or higher. Appendix IV has more information on the distribution of boys' and girls' scores.

Analysis of Year 7 reading comprehension link questions using a t-test indicated that there was no statistically significant difference between the 1992 and 1991 performances on link questions.

See Appendix II for a further description of performance on link questions.

Reading for Different Purposes Test Results

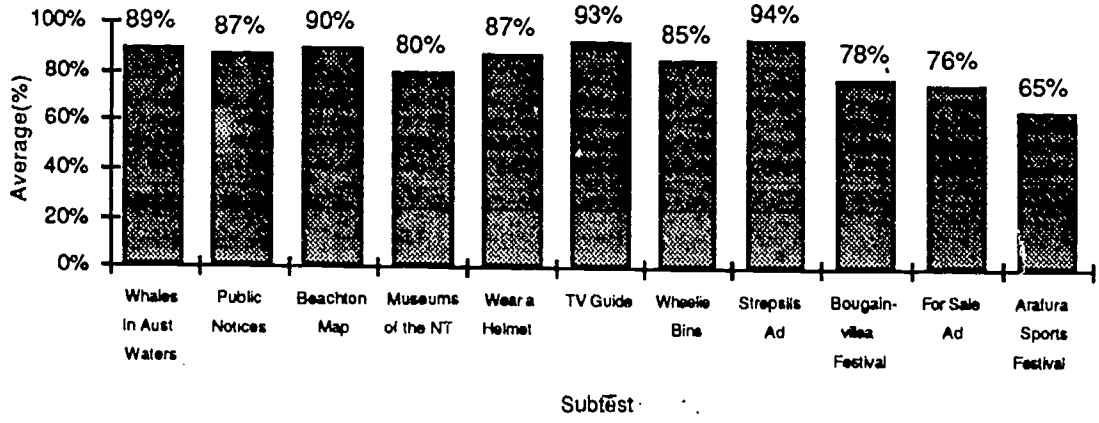
Table 8 Overall Statistics for the Year 7 RDP Test

Number of Questions	50
Average Score	42
Number of Students	1760
Standard Deviation	7.52

Fewer students sat this test than in 1991 when 1845 students sat the test. Once again there was a fair amount of variation in scores as indicated by the size of the standard deviation.

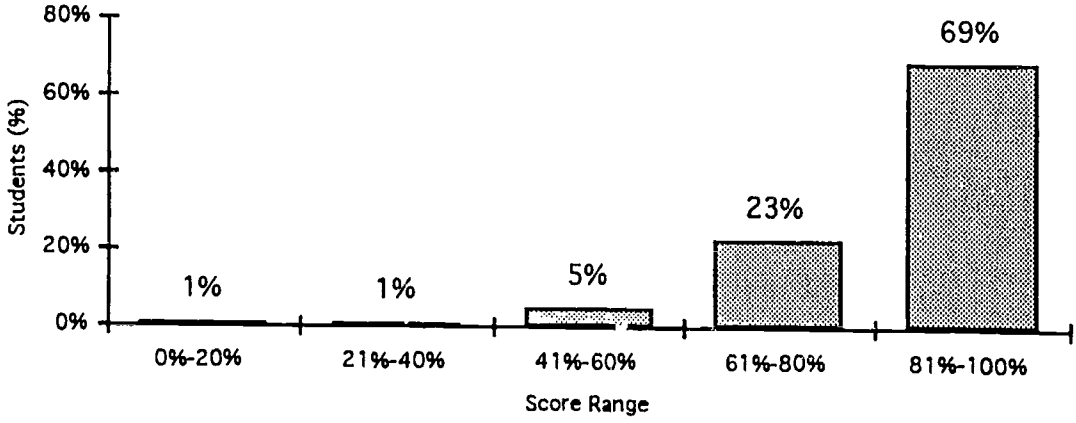
In this test forty-two students achieved the maximum possible score of fifty. All of the four students who scored zero were boys. The lowest score by a girl was eight.

Graph 25 Average Subtest Scores for the Year 7 RDP Test



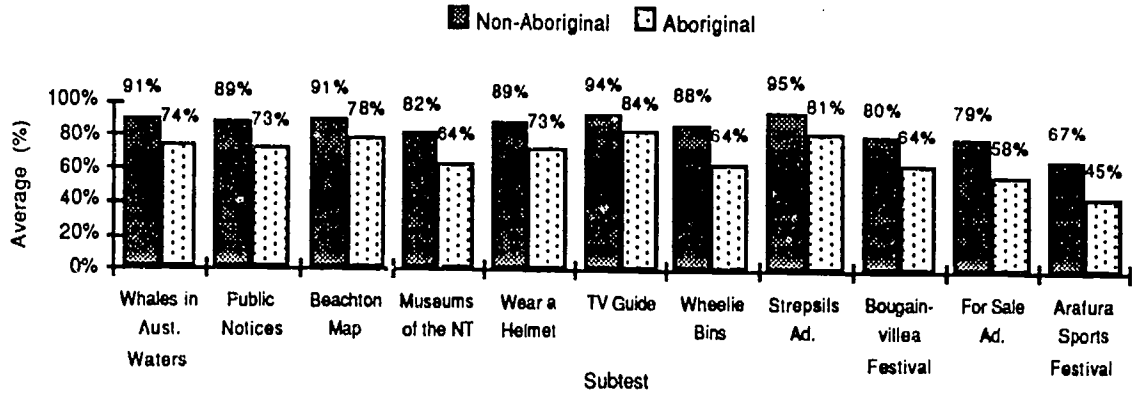
The TV Guide and the Strepsils advertisement were the easiest subtests. Students found the Arafura Sports Festival timetable relatively difficult.

Graph 26 Distribution of Scores for the Year 7 RDP Test



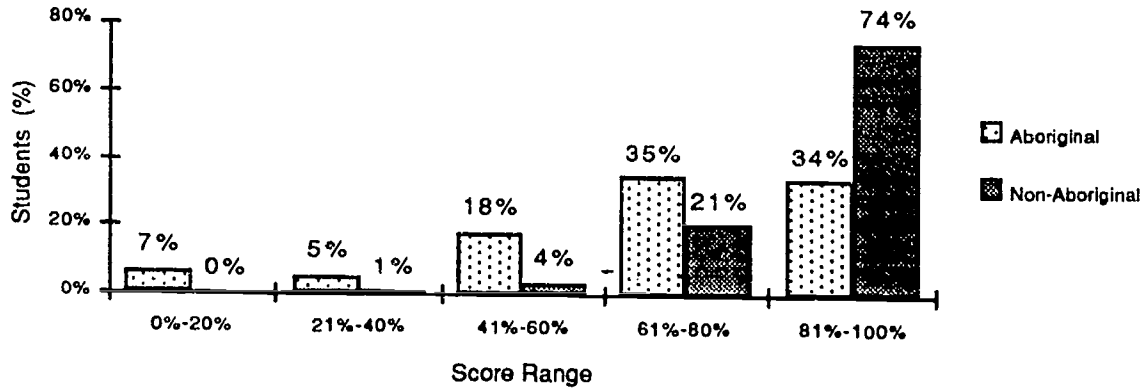
The nature of the above distribution shows that most students found this to be quite an easy test. This distribution shows that 92.39% (1627) of students scored 61% or better on this test. Only 2.38% (42) of students had scores of 40% or less.

Graph 27 **Average Scores of Non-Aboriginal and Aboriginal Students for the Year 7 RDP Test**



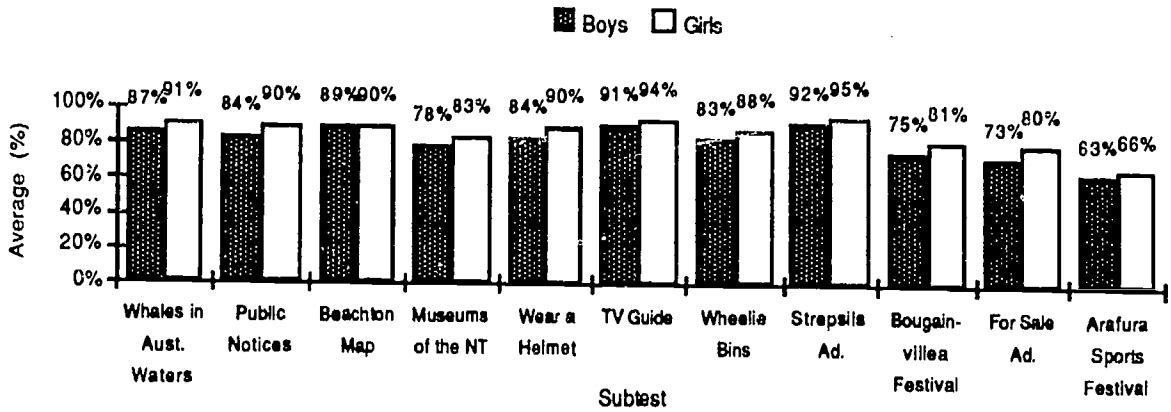
On average scores of Aboriginal students were 17% lower for each subtest. The subtests where the differences were largest were Wheelie Bins and Arafura Sports Festival. Aboriginal students' scores were closest to their non-Aboriginal peers in the TV Guide and Beachton Map subtests.

Graph 28 **Distribution of Scores of Aboriginal and Non-Aboriginal Students on the Year 7 RDP Test**



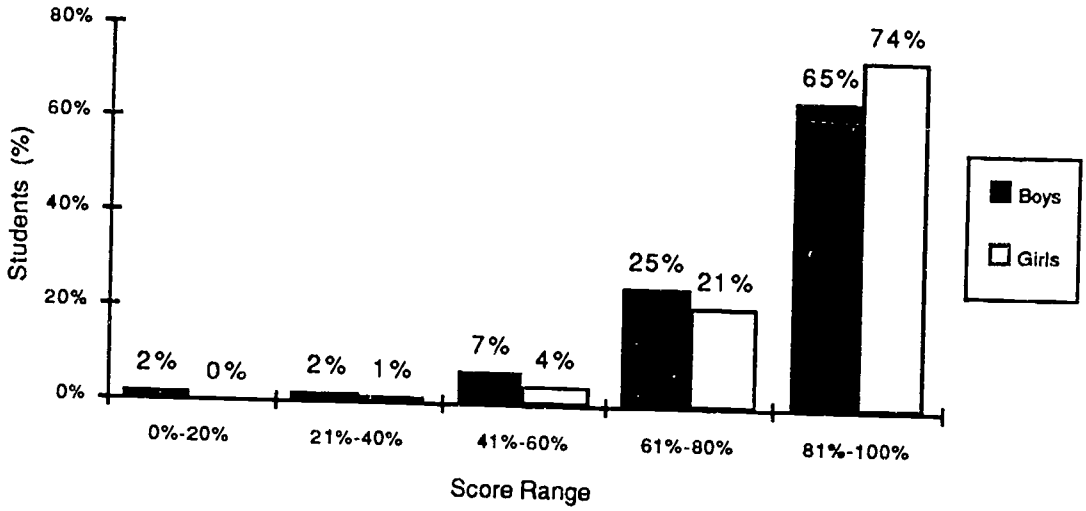
Seventy per cent of Aboriginal students who sat the Year 7 RDP test achieved a score of 61% or better. Compared with the scores of non-Aboriginal students, 26% of Aboriginal students totalled a score above the average non-Aboriginal score.

Graph 29 **Average Scores for the Year 7 RDP Test by Gender**



As with all other reading tests, girls outperformed boys. On average the difference was 4%. Girls' and boys' performance came closest to each other in the Beachton Map subtest and were furthest apart in the For Sale advertisement subtest.

Graph 30 **Distribution of Boys' and Girls' Scores on the Year 7 RDP Test**



Ninety per cent of boys who sat this test scored 61% or higher; 95% of girls were in this range. With Aboriginal students this difference was more exaggerated with 58% of boys and 78% of girls occupying this part of the scoring continuum. Refer to Appendix IV for further information on gender differences in the distribution of scores.

Scores on the link items in the Year 7 RDP test were higher in 1992 than 1991. A t-test analysis revealed that the difference was statistically significant.

Mathematics Test Results

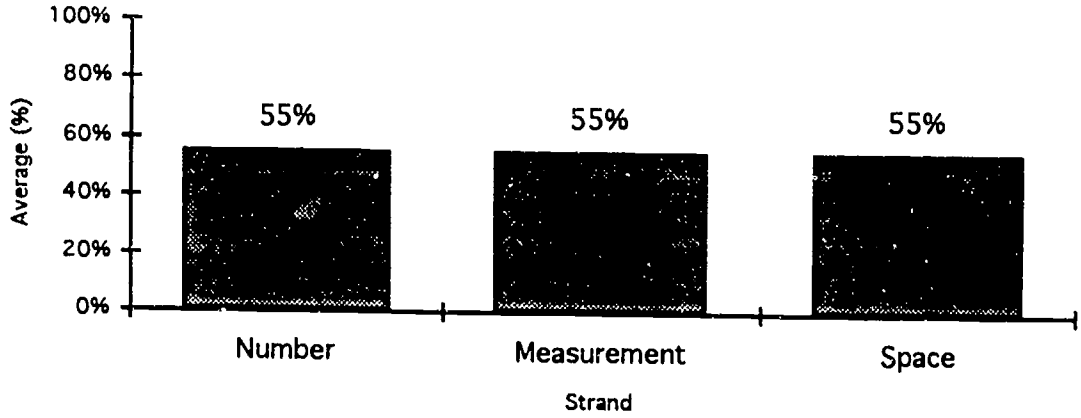
Table 9 **Overall Statistics for the Year 7 Mathematics Test**

Number of Questions	60
Average Score	33
Number of Students	1779
Standard Deviation	13.75

In 1991 1853 students did this test. The standard deviation of 13.75 was the largest for any of the PAP tests described in this report and points to a large amount of variation in scores. This points to a cohort with a most varied range of mathematical ability.

Of the five students who achieved the maximum possible score, four were girls. Four students scored zero.

Graph 31 **Average Subtest Scores for the Year 7 Mathematics Test**



Average scores for each strand indicate that students found each strand to be of equal difficulty. This result may be viewed as being somewhat unusual as in last year's test there was a range of difficulty with measurement's average per cent correct at 51%, number at 56% and then space at 59%. Measurement is often perceived as the most difficult strand but this depends a lot on the individual questions included in each subtest.

Figure 4 Year 7 Number Strand Level of Difficulty Analysis

SCORE RANGE*	QUESTION NUMBER	DESCRIPTION OF TASK
0–25%	NIL	
26–50%	48	Multiply fractions by a whole number: 1/8 x 15
	7	Problem-solving following a given rule
	58	Determine mean (average) of scores
	56	Find minimum number of picks to get a pair of socks
	45	Calculate per cent profit in a simple transaction
	20	Complete a number sentence
	21	Relationship between fractions and decimals
	54	Problem-solving using multiplication of whole numbers
	11	Round decimal to nearest tenth
	6	Multiplication of money by a whole number: $ 5.08 x 39
	31	Translate a written expression to a numeral
	39	Identify order of fractions, decimals and per cents
51–75%	46	Complete number sentence with powers of a whole number
	12	Interpret Venn diagram
	17	Identify largest from decimals, fractions and per cent
	25	Solve problem by finding remainder: 487 ÷ 9
	51	Subtract fractions then reduce to simple terms
	34	Calculate map distance, then round to nearest 100
	23	Find largest product of four multiplication algorithms
	52	Interpret pie graph
	13	Interpret line graph
	16	Calculate left over fraction in cake problem
	33	Addition of decimals: 54.08 + 35.89 + 7.3
	36	Multiply a whole number by fractions then subtract
	5	Regrouping of numbers
	42	Solve a rate of flow problem
76–100%	35	Solve a money problem: ($18.76 – $7.80)/4 kg
	2	Subtract decimals: $60.00 – $5.08
	32	Continuation of a whole number pattern
	4	Use of Roman numerals

* Score Range shows the per cent correct category into which each question falls. For example, the twelve questions in the 26–50% category were the most difficult.

Three of the five lowest scoring questions were 'problems', ultimately involving simple operations, where identifying the procedure required to solve the 'problem' appeared to cause the difficulty. Question 48, involving multiplying a whole number by a fraction, was the most difficult number question, yet one of the easier questions, 36, essentially involved multiplying a number by a fraction then subtracting.

Figure 5 Year 7 Measurement Strand Level of Difficulty Analysis

SCORE RANGE	QUESTION NUMBER	DESCRIPTION OF TASK
0–25%	NIL	
26–50%	30	Compare areas of triangles
	43	Calculate cost of sending parcel by converting kg to g
	59	Calculate area of polygon
	19	Order measurements involving grams, kilos and tonnes
	26	Calculate perimeter of right-angled polygon
	10	Calculate perimeters
	60	Calculate time elapsed between times on two clock faces
	44	Determine volume of given solid shape
51–75%	55	Ordering of measurements involving cm and m
	24	Use the relationship between triangles and rectangles
	8	Find the volume of a rectangular prism
	41	Calculate area of a parallelogram
	15	Combine mL and L to calculate volume
	47	Problem-solving: calculation based on perimeter
	1	Estimate mass of an object with volume given
	3	Draw specified time on clock face
76–100%	37	Convert 12 h time to 24 h time
	40	Solve problem involving a simple timetable

Eight (44%) of the measurement strand questions were answered correctly by between 25 and 50% of students, while only two questions were correctly answered by 76% or more students.

Calculating perimeters and areas, as well as using the relationship between units used in measuring mass, proved the most difficult operations for students attempting measurement strand questions. The three easiest questions involved understanding the measurement of time.

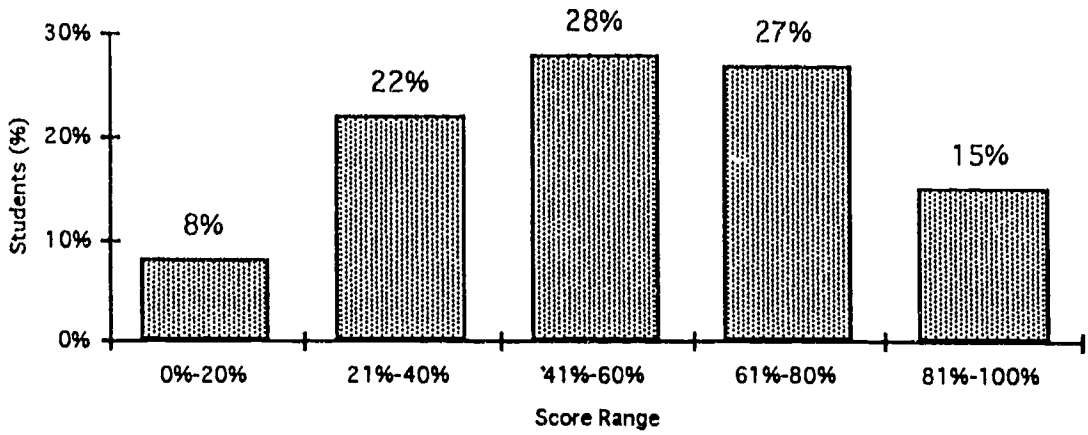
Figure 6 Year 7 Space Strand Level of Difficulty Analysis

SCORE RANGE	QUESTION NUMBER	DESCRIPTION OF TASK
0–25%	NIL	
26–50%	28	Use a scale to find a distance
	18	Identify the vertices, faces and edges of a shape
	53	Give the number of lines of symmetry in a polygon
	29	Given three angles, determine fourth angle of a quadrilateral
	57	Determine the shape of a dissection
51–75%	50	Locate points on a grid to spell out message
	49	Solve a problem involving area and congruent triangles
	14	Calculate area using scale
	38	Calculate the sum of internal angles of a triangle
	22	Find final direction after movements and turns on a grid
	9	Investigate reflectional symmetry
76–100%	27	Read compass to give ship's course

Five (42%) of the twelve space strand questions were placed in the 50% or less correct category and one (8%) of the space questions was in the 76% or better category. The question that posed the greatest difficulty involved using a scale to calculate a distance. The other two most difficult questions were identifying and counting vertices, faces and edges on a prism and giving the number of lines of symmetry in a polygon.

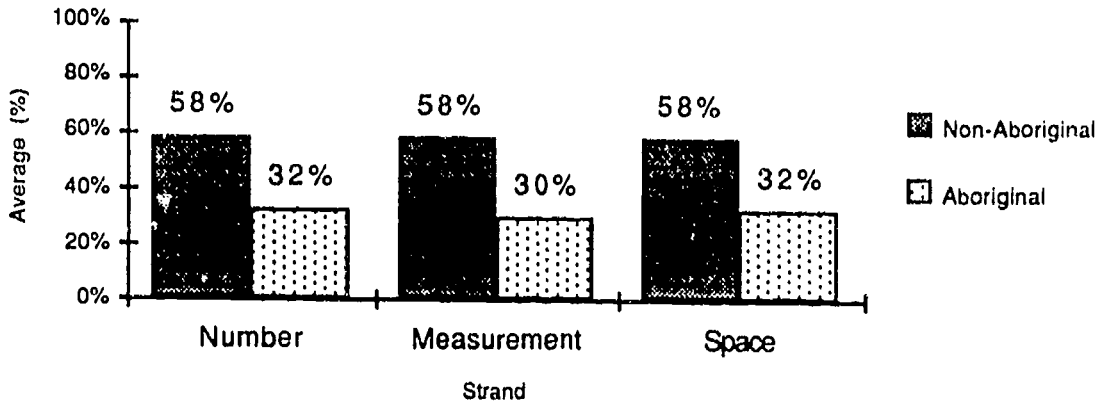
Both the easiest and hardest space questions came from the same map. Students had few problems with direction but did have some difficulty with using the scale. There were two questions involving the symmetry concept. Question 53 where the term 'symmetry' was used proved difficult; the question 9 where the word 'reflection' was used was one of the easier questions.

Graph 32 Distribution of Scores for the Year 7 Mathematics Test



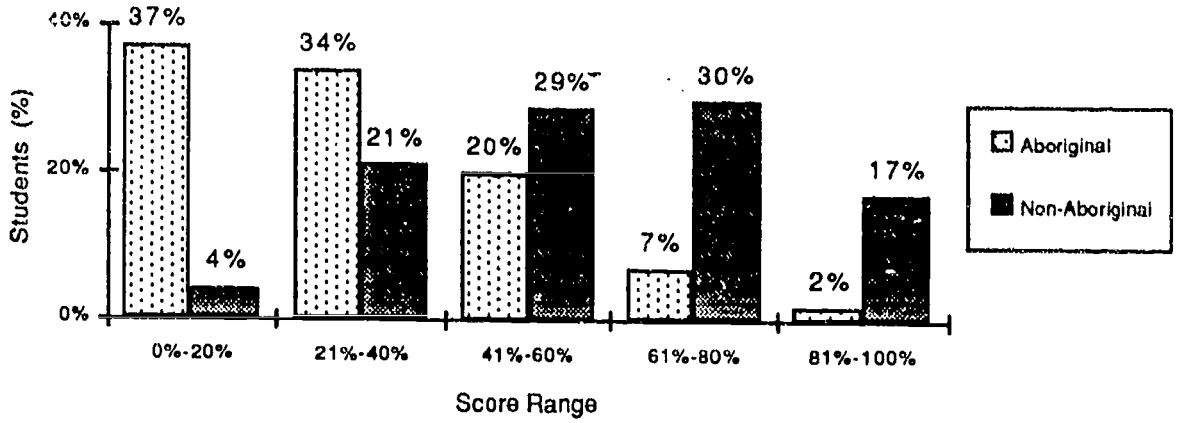
As with the distribution of Year 5 mathematics scores (see Graph 14) a shape approximating a normal distribution resulted, with the average score being towards the middle of the score range continuum and scores tapering towards both ends of this continuum. Four hundred and ninety students' scores fell within the 41% – 60% score range and this is also where the average score of 55% was.

Graph 33 **Average Scores of Non-Aboriginal and Aboriginal Students for the Year 7 Mathematics Test**



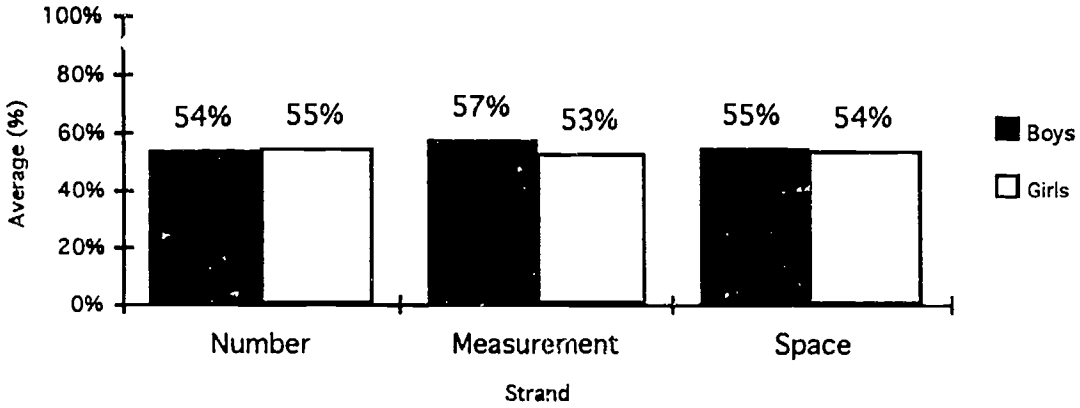
Non-Aboriginal students scored better than Aboriginal students in all subtests by an average of 27%, with the difference being marginally larger for the measurement subtest questions. On average, Aboriginal students clearly found this to be a relatively difficult test.

Graph 34 **Distribution of Scores of Aboriginal and Non-Aboriginal Students on the Year 7 Mathematics Test**



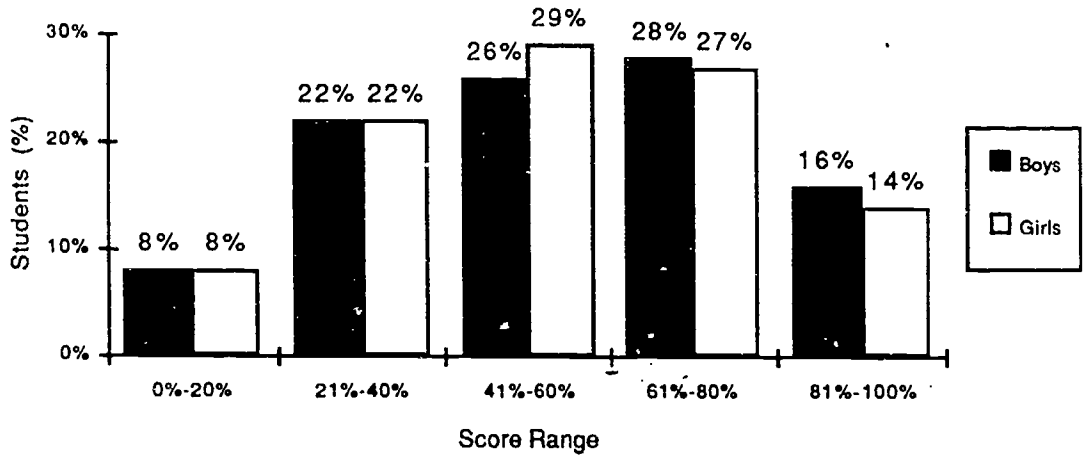
In this relatively difficult test, 9% of Aboriginal students were able to get over 60% of the test questions correct. Ten per cent of Aboriginal students who sat this test scored above the non-Aboriginal average.

Graph 35 Average Scores for the Year 7 Mathematics Test by Gender



Boy's performance was slightly better than girls. This was largely due to their higher scores in the measurement subtest. There was little variation in scores for the other two subtests.

Graph 36 Distribution of Boys' and Girls' Scores on the Year 7 Mathematics Test



Marginally higher proportions of boys than girls who sat this test achieved a score of 61% or better (44% of boys compared with 41% of girls). The opposite occured with Aboriginal students where 8% of boys were in this category against 10% of girls. See Appendix IV for a fuller account of gender differences in the distribution of scores.

Analysis of Year 7 mathematics questions using a t-test indicated that the difference between the 1992 and 1991 performances on link questions was not statistically significant.

See Appendix II for a further description of performance on link questions.

4. FUTURE DIRECTIONS: Analysing Performance in Mathematics using Item Response Theory

The following is a summary of a paper presented by ZS Huang, Principal Research Officer, to the Second Australian Council for Educational Research National Conference on Assessment in the Mathematical Sciences, at Surfers Paradise, Queensland, 15–18 April 1993.

Use of Item Response Theory (IRT) is now a common method of analysing and reporting performance on assessment tasks in most Australian states. It should be noted that the term item as used in this report is freely interchangeable with the term question. The purpose of this chapter is to foreshadow the introduction of IRT in analysis of PAP test results and to offer an explanation, supported by examples, of IRT being used in describing the performance of the system and an anonymous student from this year's Year 5 and 7 PAP mathematics tests.

Analysing Performance on Individual Questions at Year 5

One of the most useful statistics from the first report is the p-value or proportion correct for individual questions used extensively in the PAP. However, the discussion that follows is an attempt to take advantage of recent developments in Australia in bringing modern measurement theory into the classroom, and perhaps point the direction in which results can be reported for the PAP in future.

The mathematics questions used in 1992 for Years 5 and 7 have been 'calibrated' based on their estimated difficulty levels using the Rasch model in IRT (Wright and Stone 1979; Masters et al, 1990). A computer software program, *The Interactive Test Analysis System*, developed by Adams and Khoo (1992) was used to calibrate the questions and generate the diagrams shown in Figures 7, 9 and 10, pp33, 36, 38. (Adams, R and Khoo, Giek-Toon, 1992, *The Interactive Test Analysis System*, Australian Council for Educational Research, Hawthorn, Victoria).

The Rasch model enables one to place all questions on the same scale based on their estimated difficulty levels. An arbitrary scale of ten to ninety has been used for convenience to avoid the notion of a percentage score. A student's achievement on the total test or any subset may be expressed as a score on this scale continuum. The location of questions along this difficulty continuum indicates how difficult they are relative to each other. Questions located at positions higher on the scale are more difficult as they are answered correctly by a smaller proportion of students; questions at positions lower on the scale are easier as they are answered correctly by a greater proportion of students.

The result of calibrating the sixty questions done by Year 5 students in 1992 and positioning them along a scale continuum is shown in Figure 7, p33. One may call it an item response 'map' for Year 5 in which all the questions in each of the strands, namely number, measurement and space, are positioned along a continuum of

increasing difficulty from bottom to top. The questions towards the bottom end of the scale such as N31, M36 and S8 were among the easiest as these were answered correctly by most students whereas the questions towards the top end such as N6, M11 and S28 were difficult as they were answered correctly by fewer students. The letters N, M and S represent the number, measurement and space strands and the numeral after the letter shows the question number on the 1992 test. In the following analysis, the nature of some selected questions has been described in an attempt to explain their positions on this continuum and highlight the questions which appear to be typically easy or hard for Year 5 students in the Territory.

Firstly, in the number strand, the easier questions towards the bottom of the scale were those involving arithmetical computations such as N10, N31, N45 and N2. Most students could add and multiply whole numbers and compute money problems. Located near the middle of the scale and hence more difficult were questions testing the concept of simple fractions such as 3/8 = ? /16 in N54 and 1 - 3/5 = ? in N15 in which 50–60 per cent were successful. The questions became more difficult in N50 and N9 where multiples and factors were investigated. Towards the higher end were questions which many were unable to do or did not attempt. In N22, students must know what algorithm to use to determine the length of the remaining piece of rope after two pieces have been cut from it. In N24, an understanding of place values was tested and in N6 the use of fractions to express a proportion.

In the measurement strand, among the easier questions appearing near the bottom end of the scale were M36 and M5 on measuring capacities, M29 on reading a bar graph and M33 and M48 on finding the weights and volumes of cubical objects. These had relatively high rates of success. Towards the middle and hence harder were questions such as M16 on finding the area of a regular shape and a similar task in M43 on drawing a shape with an area of 13.5 square units. Towards the top of the scale were questions in which students had less success. For example, M12 on calculating from a timetable the total duration of stops on a flight from Darwin to Alice Springs in which fewer than Forty per cent were successful and M52 on finding the weight of two fish in kg given their weight in g. The most difficult question in this test was M59 that involved calculating the length of fence around a yard where the lengths of some sections had to be deduced first. Fewer than twenty per cent succeeded.

On the whole, students did well on the space strand. In S8 about eighty-five per cent correctly gave the number of triangular shapes needed to make a hexagonal shape, thus placing this question near the bottom end of the scale. Moving an object in space following compass directions in S25 was relatively easy. Slightly more difficult were questions such as S35 on drawing a shape twice as large and S57 on drawing an object following rotation. The success rate for these was about sixty-five per cent. Drawing lines of symmetry in S19 proved difficult for fifty-five per cent of students. The hardest question in the space strand was S28 on finding the number of vertices for a regular hexagonal prism.

The pattern of results displayed in Figure 7 provides a useful basis for comparing the relative difficulty of questions within and, if appropriate, across strands and also in identifying clusters of content/skill areas which require further attention. It is likened to a 'map' because one obtains a picture of the kinds of tasks which students find easy or hard from their locations along the scale. The Rasch model can also provide a useful basis for measuring student progress over time and for monitoring achievement levels from one year to the next. A few questions from the 1992 Year 5 test have been reproduced in Figure 8.

Figure 7 System Map Showing Locations of Year 5 Questions

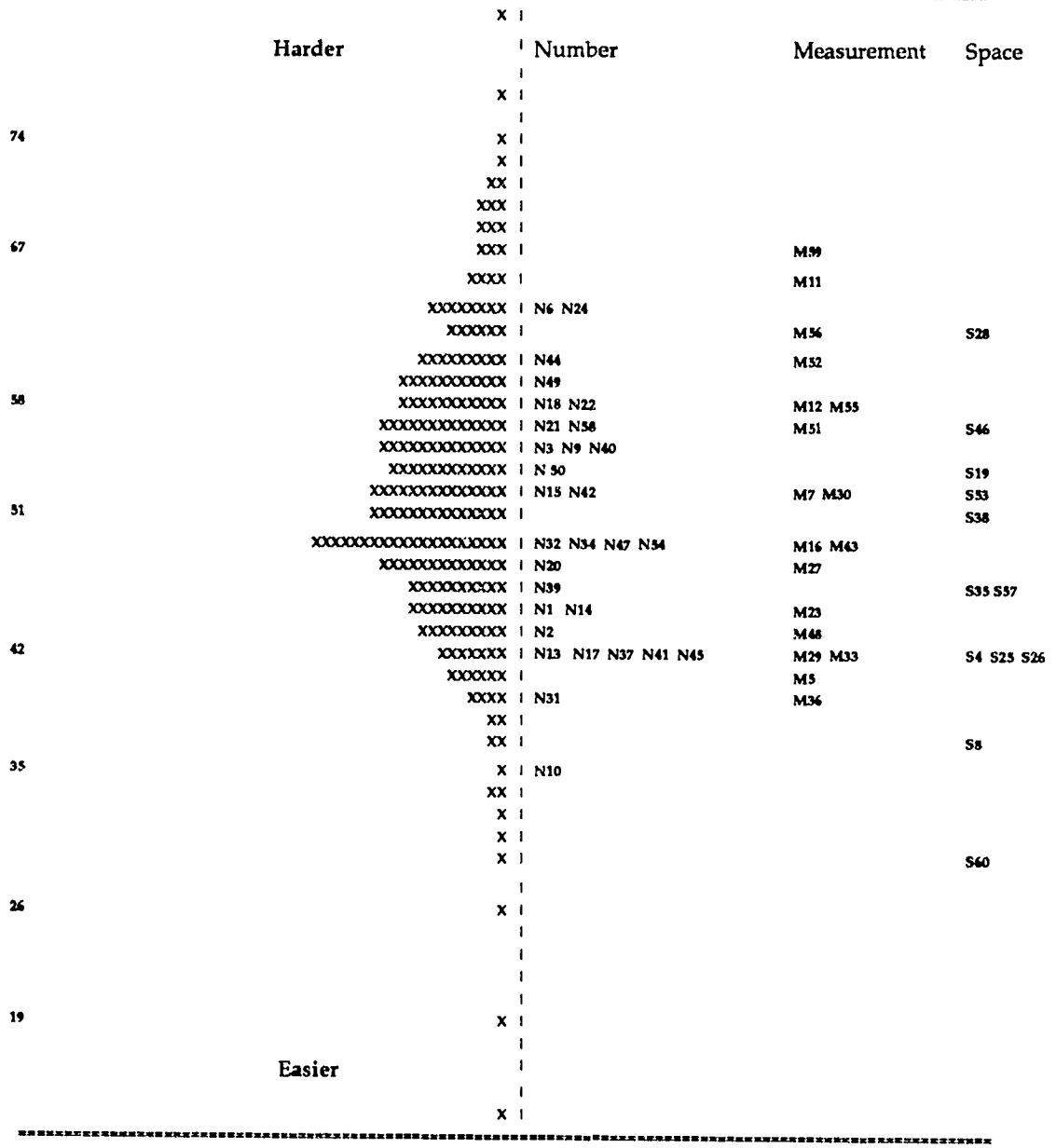
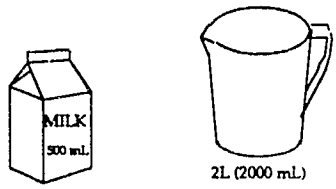


Figure 8 Sample of Year 5 Questions

M36

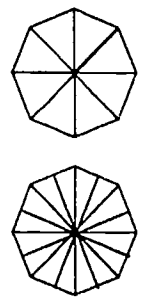


How many milk cartons will I need to fill the jug?

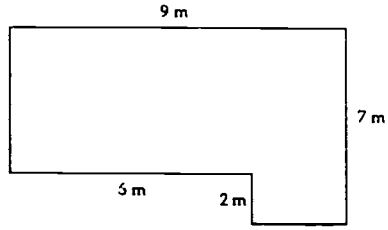
Answer: _____

N54 Complete this number sentence (you can use the drawing to help you).

$$\frac{3}{8} = \frac{\quad}{16}$$

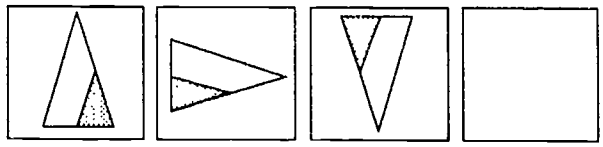


M59 What is the length of the fence around this yard?



Answer: _____

S57 Complete the pattern by filling in the blank square.



Analysing Performance of Individual Students at Year 5

A map showing the locations of all questions in a test along a continuum of difficulty is useful in identifying which questions the students in the target population find easy and which they find difficult. However, a different kind of map may be constructed to show how each individual student has performed on specific questions in the test. An individual student map is constructed in Figure 9 based on an analysis of the 1992 results of an anonymous Year 5 student. Questions answered correctly are shown on the left of the vertical scale down the centre; questions answered incorrectly are shown on the right.

As in Figure 7, the same questions are located at their estimated difficulty levels on the scale; the more difficult questions appear towards the top of the scale, and easier questions towards the bottom. The results from this analysis show that the student had thirty-six of the sixty questions correct or sixty per cent. The student's estimated achievement score on an arbitrary scale of 10 to 90 was 53.5 as shown by the crosses (xxx) on the scale. The band between the dotted lines gives an indication of the error associated with this estimate. The dotted line on the left is an estimate of the upper boundary of the student's achievement level and the dotted line on the right is an estimate of the lower boundary of his/her achievement level. These two dotted horizontal lines mark out four corners of the map.

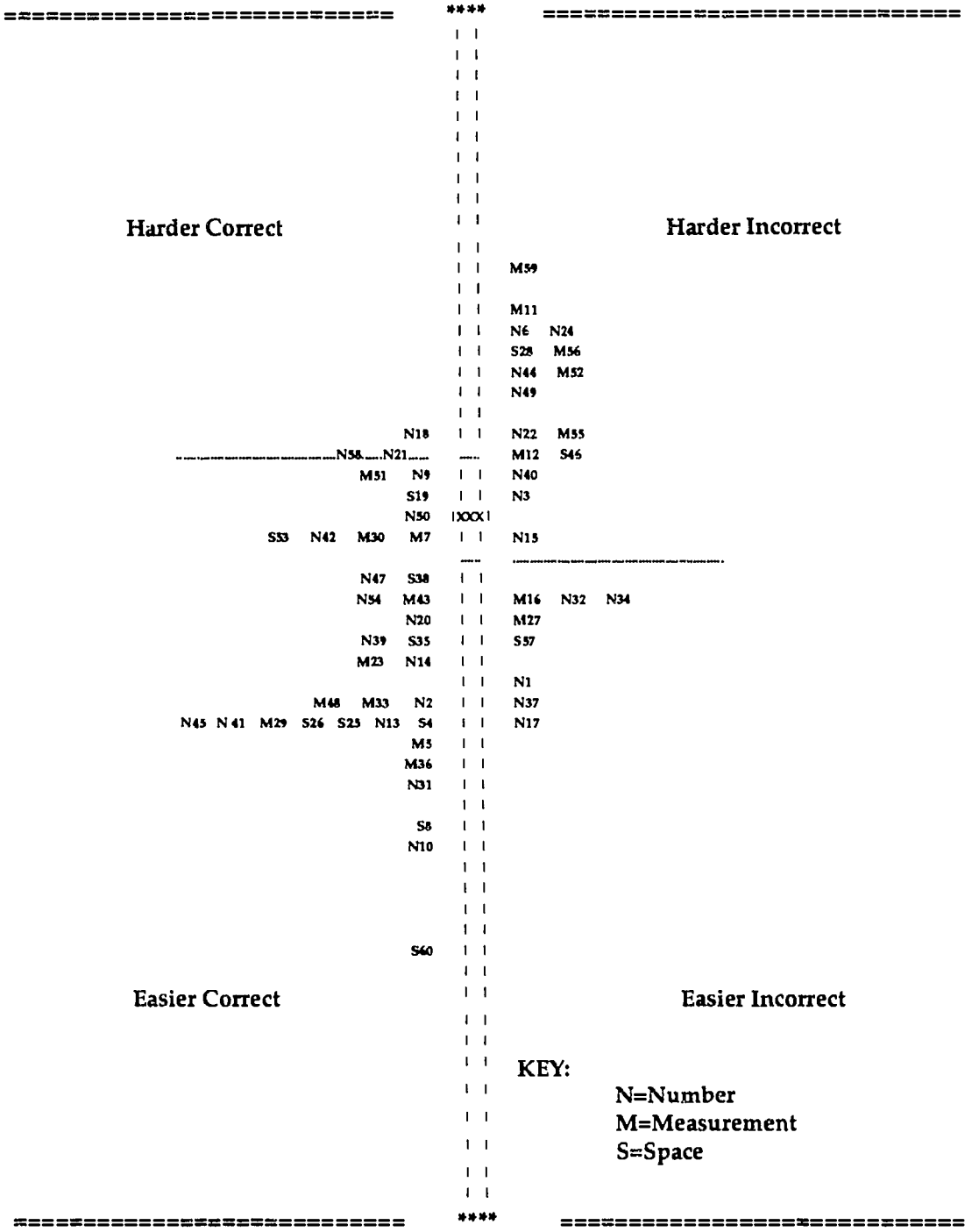
Of particular interest are questions in the top left and bottom right of the map as they reveal unexpected results. There are three questions, namely, N18, N58 and N21 in the top left corner above the line marking the student's estimated achievement level. The top left corner reveals questions which were considered too difficult for the student yet they were answered correctly. The locations of these questions also indicate that they were quite difficult for many Year 5 students. Hence, they reveal some special strengths of this particular student. These were in manipulating codes, counting (combination) and place values.

One would have expected that the student would have correctly answered questions in the bottom right corner of the map as these are below the dotted line and hence within the student's estimated achievement level. But the student was unsuccessful in these questions, namely, N17, N37, N1, S57, M27, M16, N32 and N34. An analysis of these questions shows that the student had difficulties with number and measurement concepts which do not appear to be difficult for most students. These were in multiplication (regrouping), subtraction (decimals), place values, area of polygon, time conversion (minutes into seconds and vice versa) and rotation in space.

The top right corner reveals other questions that the student did not answer correctly or did not attempt. These questions were also difficult for most students in Year 5 as evidenced by their locations on the scale. They include subtraction of fractions, place values, problem-solving, perimeter, weight conversion and time interval. An individual map allows one to analyse the pattern of results for each student and report on each student's strengths and weaknesses. This analysis is based on a knowledge of the student's estimated achievement level on the whole test and the positions of questions calibrated along a scale of difficulty.

Figure 9 Individual Student Map

School Code: 1 % score: 60.0
Student Number: 003 PAP score: 53.5



Analysis of Results for Year 7

Figure 10, p38, shows the pattern of responses to questions (items) in the 1992 Year 7 mathematics test. Two of the most difficult questions in the whole test were in the number strand, namely, N7 and N56, located in the upper part of the scale. In N7, students were asked to find the number of days required to save $30 if an initial amount of $1 is doubled on each consecutive day. Question N56 was a problem-solving task involving chance and counting. A percentage problem in N45, and a fraction task in N21 were less difficult but still posed a problem for the majority.

Located near the middle of the scale is N51 requiring students to subtract one fraction from another (same denominator) and reduce the answer to its simplest terms. About half of Year 7 students gave the correct answer. Two other questions of interest, located near the centre were N39 and N17, which tested mastery of proportions expressed as fractions, decimals and percentages. Slightly lower on the scale was question N34, which involved adding distances from a map and rounding the answer to the nearest 100 km.

Towards the lower end of the scale, in N52, seventy per cent of students had no difficulty in interpreting percentages from a pie chart, and in N13, in finding the cost of hiring an aircraft for a given distance from a line graph. Money problems were easy for most as in N35 and N36.

In the measurement strand, the most difficult questions were M30, M43 and M59 but still easier than four number questions located at the higher end of the continuum. These involved finding out which two triangles within the rectangles have the same area, in M30; calculating postage charges from the variable rates specified, in M43; and finding the area of a regular polygon where some measurements were not given, in M59.

Slightly less difficult was a task in M19 of numerically ordering weights expressed in grams, kilograms and tonnes. Nearly fifty per cent could work out the distance covered by two laps around a rectangular field measuring 160 m by 140 m, in M10. Fifty per cent had difficulty in M60 in calculating the time lapse between two times shown on a clock face. Conversion of metric lengths, in M55, was done correctly by slightly over half of all students.

One of the easiest questions (M15) was adding the capacities, expressed in millilitres and litres, of two containers and then converting the answer to litres. Most students could show time on a clock face or digital clock in M3 and M37.

Year 7 students did not find the space questions difficult on the whole. One of the more difficult questions was S53 on lines of symmetry, located above the centre. Nearly fifty per cent could, in S57, identify the shape of the cross-section of a cut made diagonally on a cuboid.

Nearly sixty per cent could correctly calculate area on a map using a scale, in S14. Most knew the angle properties of a triangle, in S38 and were able to follow compass directions, in S22 and S27, at the lower end of the continuum.

Some questions from the Year 7 test are reproduced in Figure 11, p39.

Figure 10 System Map Showing Locations of Year 7 Questions

Scale				
	XX			
	Harder	Number	Measurement	Space
81				
	XX			
	XXX			
73				
	XXX			
	XXX			
	XXXXX			
	XXXX			
64	XXXXXXXXX			
	XXXXXXXXXXXXX			
	XXXXXXXX			
	XXXXXXXX	N7 N48		
	XXXXXXXXXXXXXXXX	N56 N58		
	XXXXXXXXXXXXXXXX			
55	XXXXXXX	N45	M30 M43 M59	
	XXXXXXXXXXX	N20 N21		S28
	XXXXXXXXXXXXX		M19	S18 S53
	XXXXXXXXXXXXXXXX		M26	
	XXXXXXXXXXXXXXXXXX	N6 N11 N31 N54	M10 M44 M60	S29 S57
	XXXXXXXXXXXXXXX	N39	M35	S49 S50
	XXXXXXXXXXXXXXXXXXX	N12 N17 N25 N46 N51	M24	
45	XXXXXXXXXXXXXXXX		M8 M41	S14
	XXXXXXXXXXXXXXX	N23 N34		S38
	XXXXXXXXXXXXXXXX		M15	
	XXXXXXXXXXXXXXX	N13 N16 N52		S22
	XXXXXXXXXXXXX	N33 N36	M47	S9
	XXXXXX			
36	XXXXXXXXX	N5 N42	M1 M3	
	XXXXXXXXXXXXX	N2 N35		
	XXXXX	N4 N32		
	XXXXXXX		M37 M40	
	XXX			
	XXX			
27	XXXXX			S27
	XX			
	XXX			
	XX			
	XX			
	Easier			
19	XX			
	X			

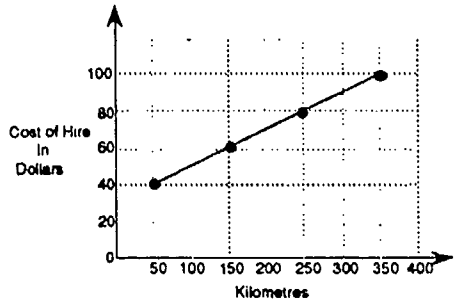
Each X represents 5 students N=1779 L=60 Mean=48.05

Figure 11 Sample of Year 7 Questions

N7. A boy is offered a job which pays $1 on the first day. His pay doubles on each day after this. How many days will he have to work in order to save $30?

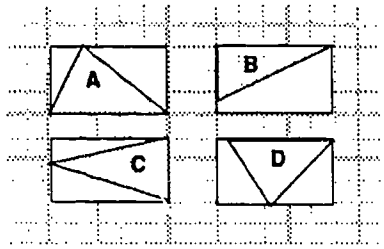
Answer: _____

N13. The graph shows the cost of hiring an aircraft for different distances. Find the cost of a 200 km trip.



Answer: $ _____

M30. Which two triangles have the same area?

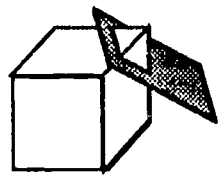


Answer: _____

M43. POSTAGE CHARGES
$5.00 for first 250 g.
$4.00 for each additional 250 g
Calculate the cost of sending a parcel weighing 1.25 kg.

Answer: $ _____

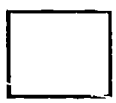
S57.



Which of these shows the face of this cut?

A) 

B) 

C) 

D) 

APPENDIX I

1992 Committees for the Primary Assessment Program

NT Board of Studies – Primary Assessment Committee (1993)

Cliff Fowler	(Chair) Chief Assessor – ERA
Ron Abbott	Principal – Driver Primary School
Nick Cockshutt	Director – Curriculum
Pascale Dixon	(Executive Officer) Senior Education Officer
Robyn Ferguson	Shepherdson College
Huang Zheng Sen	Principal Research Officer
Mick Myers	Principal – Moil Primary School
Chris Ross	Feppi

Test Construction Panels Responsible for Constructing the 1992 Tests

Year 5 English Tests

Gerry McCue	Curriculum and Assessment (Convenor)
Pascale Dixon	Curriculum and Assessment
Gaetano Gugliotta	Nakara Primary School
Julie Wills	Sanderson Primary School
Karen Taylor	Ludmilla Primary School
Bruce Francis	Berry Springs Primary School
Lisa Day	Parap Primary School
Debbie Steele	Moulden Park Primary School
Maria Mendoza	Administrative Support

Year 5 Mathematics Test

Dave Watkins	Curriculum and Assessment (Convenor)
Lakshan Badlu	Karama Primary School
Liz Monro	Moulden Park Primary School
John Moulds	Ludmilla Primary School
Peter McPhee	Anula Primary School
Robyn Blake	Wanguri Primary School
Sue Ray	Parap Primary School
Otto Eykman	Stuart Park Primary School
Geoff Gillman	Holy Spirit Primary School

Year 7 English Tests

Robbie Dunbar	Curriculum and Assessment (Convenor)
Pascale Dixon	Curriculum and Assessment
Debbie Ford	Sanderson Primary School
Kathy Banks	Ludmilla Primary School
Linda Brodie	Berry Springs Primary School
John Hubbard	Nakara Primary School
Marty Heysen	Batchelor Area School
Mali Whittaker	Humpty Doo Primary School
Donna Dawson	Administrative Support

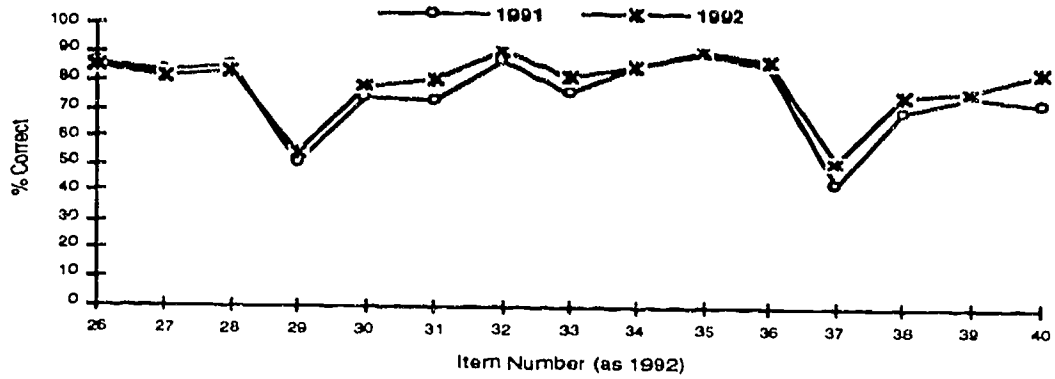
Year 7 Mathematics Test

Dave Watkins	Curriculum and Assessment (Convenor)
Anne Mauger	Moil Primary School
Josie Roberts	Moil Primary School
Fathma Mir	Parap Primary School
Peter McPhee	Anula Primary School
Bill Armstrong	Ludmilla Primary School
Leslie Hassett	Malak Primary School
Malcolm Taylor	Wanguri Primary School
Dave Henry	Nakara Primary School

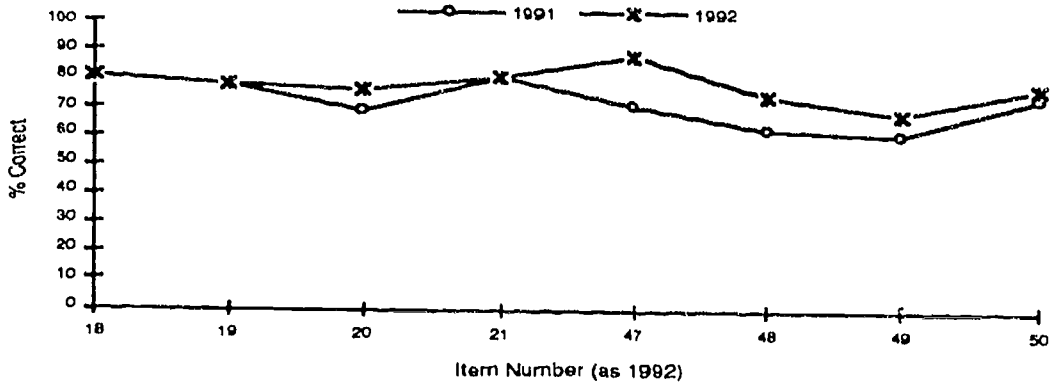
APPENDIX II

Graphic Representation of 1991 and 1992 Performance on Link Questions

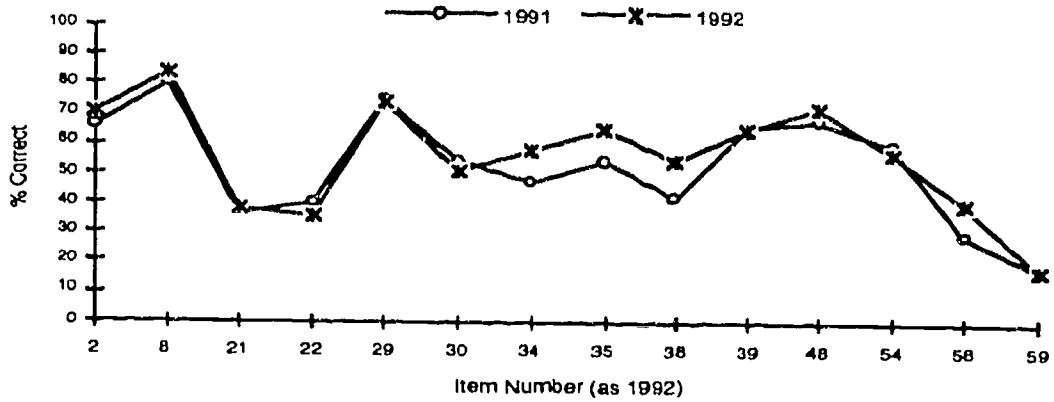
Graph 37 Comparison of Average Scores for the Link Questions Common to the 1991 and 1992 Year 5 Reading Comprehension Tests



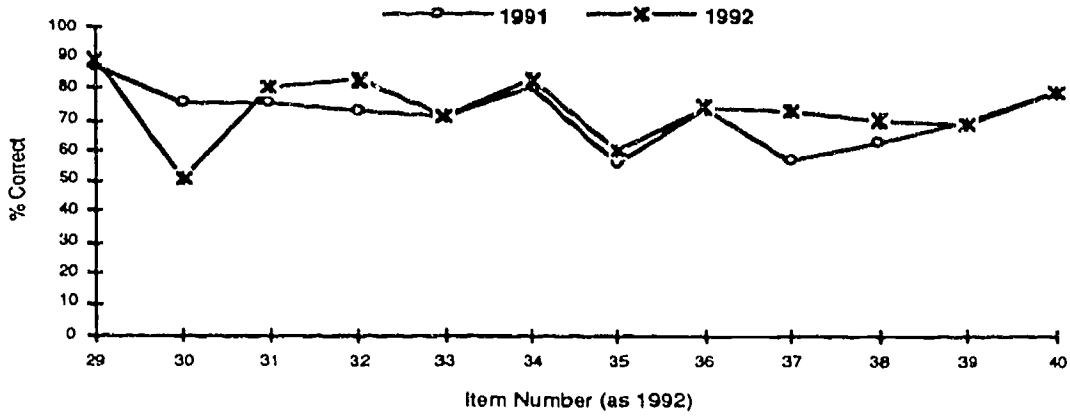
Graph 38 Comparison of Average Scores for the Link Questions Common to the 1991 and 1992 Year 5 RDP Tests



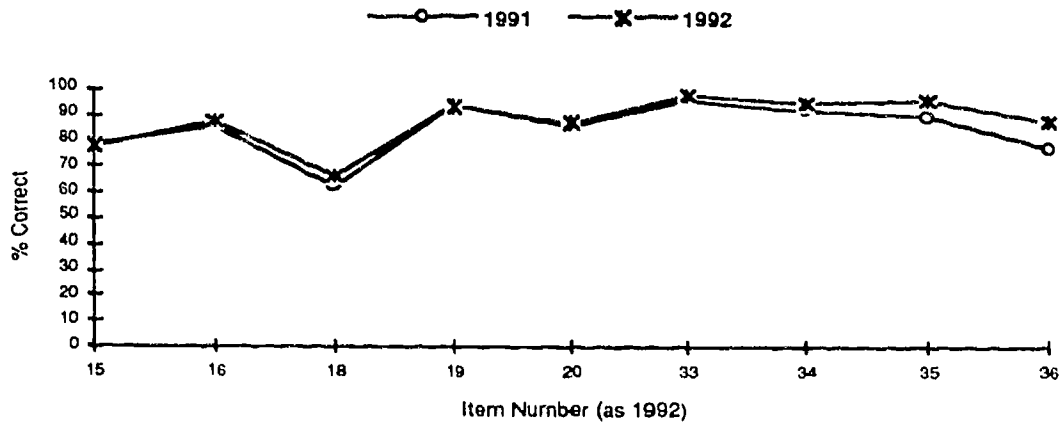
Graph 39 Comparison of Average Scores for the Link Questions Common to the 1991 and 1992 Year 5 Mathematics Tests



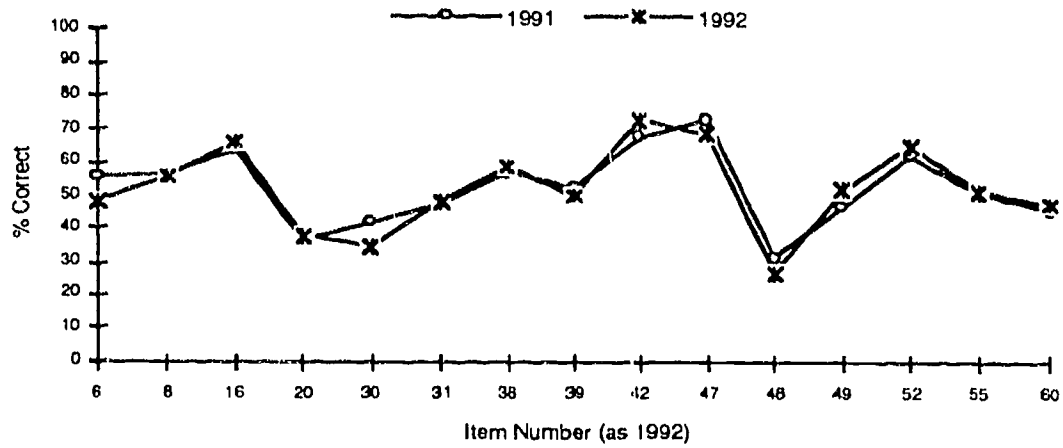
Graph 40 Comparison of Average Scores for the Link Questions Common to the 1991 and 1992 Year 7 Reading Comprehension Tests



Graph 41 Comparison of Average Scores for the Link Questions Common to the 1991 and 1992 Year 7 RDP Tests



Graph 42 Comparison of Average Scores for the Link Questions Common to the 1991 and 1992 Year 7 Mathematics Tests



APPENDIX III

Comparative Performance of Urban Aboriginal Students on 1991 and 1992 Link Questions

In the following table the performances of Aboriginal students on the link questions included in the 1991 and 1992 PAP tests are analysed. In the mathematics tests there was a selection of fourteen link questions. Link questions in the reading comprehension tests included a complete text and related questions. Two pieces of stimulus material and related questions made up the links in each RDP test.

Some care should be taken with the following figures as the classification of students as Aboriginal was done by the process of self-identification which means that the number of students identified as being Aboriginal may or may not reflect the true numbers of urban Aboriginal students at Years 5 and 7.

Table 10 **Performance on 1991 and 1992 Link Questions by Urban Aboriginal Students**

TEST	1991 MEAN (%)	1992 MEAN (%)	DEGREE of FREEDOM	t-VALUE	t-TEST CRITICAL VALUE	RESULT
Year 5 reading comprehension	59.41	52.36	28	-4.55	2.048	Significantly different
Year 7 reading comprehension	52.16	45.36	22	-2.68	2.074	Significantly different
Year 5 RDP	52.70	50.19	14	-0.90	2.145	Not significantly different
Year 7 RDP	69.15	59.33	16	-6.27	2.120	Significantly different
Year 5 mathematics	34.37	31.35	26	-2.09	2.056	Significantly different
Year 7 Mathematics	29.57	23.33	28	-5.20	2.048	Significantly different

Statistically speaking, i.e. using a t-Test, performance declined significantly on five of the six tests, namely the Years 5 and 7 reading comprehension tests and mathematics tests and the Year 7 RDP test. Performance remained static on the remaining test which was the Year 5 RDP.

Similar numbers of self-identified Aboriginal students completed the test in each of the two years.

APPENDIX IV

Distribution of Scores by Gender for the System and for Aboriginal Students

Table 11 Distribution of Year 5 Reading Comprehension, RDP and Mathematics Test Scores by Gender

Score Range	Reading Comprehension		RDP		Mathematics	
	Male (%)	Female (%)	Male (%)	Female (%)	Male (%)	Female (%)
0%–20%	1	1	3	1	5	4
21%–40%	5	4	6	3	19	20
41%–60%	12	8	12	8	35	38
61%–80%	31	32	28	27	31	29
81%–100%	50	56	51	60	11	9

Table 12 Distribution of Year 7 Reading Comprehension, RDP and Mathematics Test Scores by Gender

Score Range	Reading Comprehension		RDP		Mathematics	
	Male (%)	Female (%)	Male (%)	Female (%)	Male (%)	Female (%)
0%–20%	2	1	2	0	8	8
21%–40%	8	7	2	1	22	22
41%–60%	20	16	7	4	26	29
61%–80%	36	34	25	21	28	27
81%–100%	34	42	65	74	16	14

Table 13 Distribution of Year 5 Reading Comprehension, RDP and Mathematics Test Scores of Aboriginal Students by Gender

Score Range	Reading Comprehension		RDP		Mathematics	
	Male (%)	Female (%)	Male (%)	Female (%)	Male (%)	Female (%)
0%–20%	7	9	17	10	25	23
21%–40%	17	13	16	10	36	33
41%–60%	23	14	18	18	31	29
61%–80%	34	33	26	24	8	13
81%–100%	19	31	24	38	0	3

Table 14 Distribution of Year 7 Reading Comprehension, RDP and Mathematics Test Scores of Aboriginal Students by Gender

Score Range	Reading Comprehension		RDP		Mathematics	
	Male (%)	Female (%)	Male (%)	Female (%)	Male (%)	Female (%)
0%–20%	13	6	14	2	36	37
21%–40%	29	23	6	4	37	32
41%–60%	23	31	21	16	19	22
61%–80%	27	26	33	36	7	7
81%–100%	8	15	25	42	1	3

APPENDIX V

Sample Year 7 RDP Text and Questions

PUBLIC NOTICES

GYMNASTICS CLASSES COMMENCE TUESDAY 28TH JANUARY, 1992

Fully equipped facility - Qualified Coaches
Reasonable fees
PCYC Berrimah and Nightcliff
Classes for Beginners to Elite

SATURDAY

9.00 - 9.45 Kindergym A (2 - 4 years)
9.45 - 10.45 Kindergym B (4 - 5 years)
10.45 - 11.45 Beginners (6 years and up)
11.45 - 1.15 Levels 1 - 3

MONDAY

5.00 - 6.30 Levels 1 - 3

WEDNESDAY

4.00 - 5.00 Beginners
5.00 - 6.30 Levels 1 - 3

For beginners levels contact Trish 27 9673
For advanced levels contact Polly 85 6791 (a/h)

Registration and Fee Paying Day
Saturday 25th January 9.00am - 3.00pm
Berrimah Gymnasium, McMillans Road (opp. C.S.I.R.O.)

NORTHERN TERRITORY
BALLROOM DANCING
ASSOCIATION

PHONE 27 6331
A/H FOR DETAILS

THE N.T.B.D.A. HOLD A
SOCIAL DANCE
EVERY 2nd and 4th SATURDAY of each month at the airconditioned
ITALIAN CLUB
EVERYBODY WELCOME!

DANCE ENROLMENTS

SATURDAY 25 JANUARY
9 AM—2 PM

•JAZZ •MODERN •CLASSICAL •TAP

Classes for adults and children from 3 years, including examination syllabus.
Phone 85 9454

MYRIA BELL TALENT SCHOOL – NIGHTCLIFF

BALLROOM dance tuition

3 pm Saturdays children's classes

Rapid Creek Ballroom Academy

Phone 85 2861

Year 7 CHRONICLE Page 2

Read the PUBLIC NOTICES and answer questions 5 to 10.

5. Gymnastic classes commence on

   A. Saturday 25 January.
   B. Tuesday 28 January.
   C. Saturday 28 January.
   D. Monday 27 January.

6. Myria Bell Talent School takes enrolments from what age?

   Ans: _____

7. Police and Citizens Youth Club NT run beginners gymnastics classes at what times?

   A. Wednesday 4.00 - 5.00
   B. Saturday 11.45 - 1.15
   C. Monday 5.00 - 6.30
   D. Saturday 9.00 - 9.45

8. Polly is the contact person for which activity?

A. Tap dance classes - adults
B. Gymnastic classes - beginners
C. Ballroom dance tuition - all ages
D. Gymnastic classes - advanced levels

9. The Registration and Fee Paying day and time for gymnastics is

A. Saturday 25 January 9.00 am - 3.00 pm.
B. Saturday 25 January 9.00 am - 5.00 pm.
C. Saturday 28 January 9.00 am - 3.00 pm.
D. Tuesday 28 January 9.00 am - 5.00 pm.

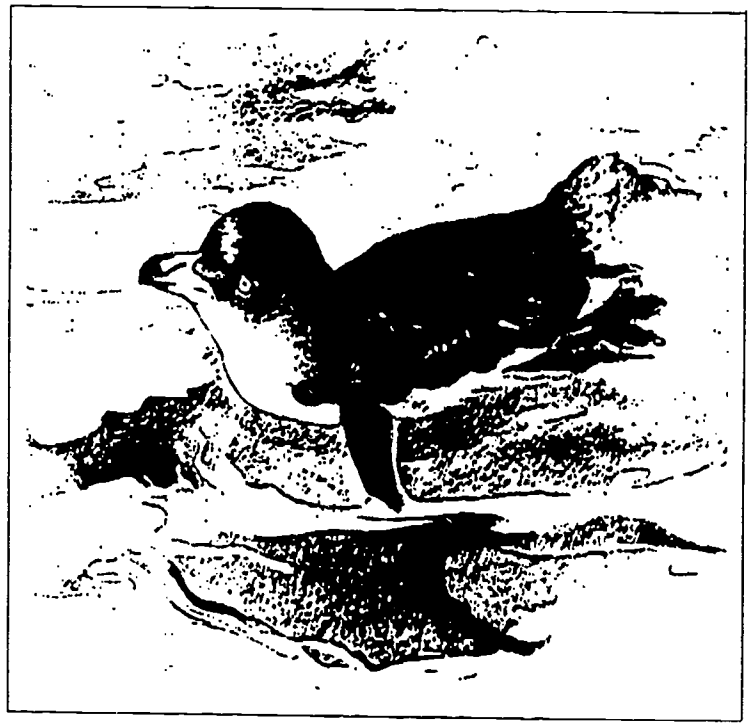
10. Write in full what the letters N.T.B.D.A. stand for.

Ans: _____

_____

APPENDIX VI

Sample Year 5 Reading Comprehension Text and Questions



Pinquo was a penguin. His home was a rocky burrow near the entrance to Sickle Bay. Hundreds of his friends lived there too, in hollows along the coast or in holes dug into the sedge-covered sandhill near by. He weighed exactly one kilogram and when he stood up he was thirty-three centimetres high. That meant that he was the smallest kind of penguin in the world. Scientists called him the Little Blue, but ordinary people said he was a Fairy Penguin. That was a beautiful and proper name for someone like Pinquo.

Pinquo was a sort of poem. His eyes were like little circles of moonlight and his feathers were soft and strong. On his back they were the colour of blue steel but his chest was as clean and white as laundered linen. At daybreak each morning he went down to the sea to fish. He swam and skipped and dived faster than the blink of an eye. He darted and swung and looped this way and that like a bird swooping about in the high clear sky. But Pinquo was flying under the water.

Pinquo was a clown. At dusk each evening he waddled up the beach towards his burrow like a small plump gentleman. He stood preening himself very particularly for a while; then he moved his flippers up and down slowly and started to sing – a dreadful song that sounded like a donkey in pain. At sea he clowned with his friends too, flipping water about or porpoising along like a skipping rocket – under the surface, up into the air, and back under again, for the sheer joy and speed of it.

Pinquo was a lovely, gentle, wonderful creature.

Read the text and answer these questions.

13. What is the popular name for the Little Blue Penguin?

Ans: _____

14. When did Pinquo go fishing?

A. Early morning
B. Late morning
C. Late afternoon
D. Late at night

15. What do Fairy Penguins eat?

Ans: _____

16. Pinquo lived in a

A. sea cave.
B. sandy hollow.
C. grassy nest.
D. rocky burrow.

17. Why was Fairy Penguin a proper name for Pinquo?

A. His feathers were soft and strong
B. He was the smallest kind of penguin in the world
C. His eyes were small, round and shiny
D. He was a very fast swimmer

18. The whiteness of Pinquo's chest was compared to

Ans: _____

19. Pinquo was compared to a bird when he

A. preened himself.
B. flapped his flippers.
C. swam under water.
D. played with his friends.

20. Pinquo was like an animal in pain when he

A. waddled.
B. sang.
C. preened.
D. dived.

21. The word **'porpoising'** suggests that Pinquo swam

A. round and round.
B. from side to side.
C. under the water.
D. in and out of the water.

22. Which word in the text **best** describes how Pinquo walked on land?

Ans: _____

23. At what time of day do Fairy Penguins return to their burrows?

A. At dusk
B. In the afternoon
C. At midday
D. At night

24. Which sentence sums up the author's feelings about Pinquo?

Ans: _____

_____

_____

25. The **best** title for this text is

A. Penguins.
B. Little Blue Penguins.
C. Pinquo the Penguin.
D. Fairy Penguins.

GLOSSARY

ABORIGINAL	Students who, on the test cover-sheet, identified themselves as being Aboriginal.
AGGREGATED	When all students' scores are totalled and reported as a system score.
ALGORITHM	Where the procedure is supplied and the student must give an answer to a mathematical problem, e.g. 5 x 3.2.
CALIBRATE	To place on a graduated scale.
CONTINUUM	The range of possible scores from zero to the maximum score.
DIAGNOSTIC	Test information that highlights areas of strength and weakness for the individual, school or system.
DISCRIMINATION	An index showing the difference in proportion of students in the upper scoring and lower scoring groups who get a test question correct.
DISTRIBUTION (of scores)	Where all students' scores are placed on a scale continuum.
ENVIRONMENTAL TEXT	Informational materials that are commonly found in students' immediate environments.
INDIVIDUAL RESULTS	An account of how each individual student performed on each test.
INFERENTIAL QUESTION	A question where the answer is not explicitly stated but can be inferred from the text.
INTERPRETIVE QUESTION	Where the answer is explicitly stated in the text and the student has to interpret the text to get the answer.
ITEM	This term is interchangeable with question
LEVEL OF DIFFICULTY	How difficult the question was as indicated by what percentage of students got it right.
LINK QUESTIONS	Common questions placed in tests over succeeding years that allow direct year-to-year comparisons to be made of system performance.
NON-ABORIGINAL	Students who, on their test cover sheets, did not identify themselves as Aboriginal.

PERCENTILE	When all scores are allocated percentiles with the highest score(s) at the one hundredth percentile. All lower scores are then allocated the remaining percentile ranks in descending order. More than one person can share a percentile rank.
PERFORMANCE	How a student or group of students did on a test expressed in terms of average score.
RELIABILITY	The measure of the amount of trust that can be placed in a test measuring the same thing, such as reading comprehension, accurately and consistently.
SCHOOL RESULTS	The combined results of all students in the school who took a test.
SIGNIFICANCE DIFFERENCE	When the difference between two results is proven, by a statistical test, e.g. a t-Test, to be unlikely to be due to chance.
STANDARD DEVIATION	The average amount in which all scores deviate from the average score.
SUBTEST	A section of the whole test that is made up of questions from the same text or strand.
SYSTEM	All of the students in the NT who have taken a particular test.
t-TEST	A statistical test that determines if a difference is significant or not.
TEXT	The written material on which a reading test is based.
VALIDITY	The extent to which a test does test what it is designed to test such as mathematical ability at Year 5.